

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	4344	100.0	4344	4	US-08-843-572E-1	Sequence 1, Appli
2	1527.6	35.2	2254	4	US-08-843-572E-3	Sequence 3, Appli
3	90.4	2.1	666	3	US-09-202-712-6	Sequence 6, Appli
4	83.2	1.9	2526	3	US-09-202-712-1	Sequence 1, Appli
5	54.2	1.2	7218	1	US-08-232-463-14	Sequence 14, Appli
6	51.4	1.2	7218	1	US-08-232-463-14	Sequence 14, Appli
7	51.2	1.2	19124	2	US-08-487-826B-13	Sequence 13, Appli
8	46.2	1.1	447	3	US-09-202-712-19	Sequence 19, Appli
9	45.6	1.0	1274	4	US-09-523-899A-1	Sequence 1, Appli
10	45.4	1.0	376	2	US-08-623-906A-18	Sequence 18, Appli
11	45.4	1.0	5852	1	US-07-867-106-2	Sequence 2, Appli
12	45	1.0	213	4	US-09-395-604A-20	Sequence 20, Appli
13	44.4	1.0	543	6	5273901-6	Patent No. 5273901
14	43.2	1.0	658	3	US-08-998-416-595	Sequence 595, App
15	42.2	1.0	4380	1	US-07-582-945-1	Sequence 1, Appli
16	42.2	1.0	4380	2	US-08-453-141-1	Sequence 1, Appli
17	42.2	1.0	4380	3	US-08-293-314-1	Sequence 1, Appli
18	42	1.0	4970	1	US-08-764-100-14	Sequence 14, Appli
19	42	1.0	4970	1	US-08-764-100-14	Sequence 20, Appli
20	41.8	1.0	12597	4	US-09-705-299-12	Sequence 12, Appli
21	41.6	1.0	10640	4	US-09-417-485D-5	Sequence 5, Appli
22	41	0.9	240	1	US-08-628-417-6	Sequence 6, Appli
23	41	0.9	397	3	US-09-253-691-3	Sequence 3, Appli
24	40.8	0.9	1037	4	US-09-181-585-3	Sequence 3, Appli
25	40.8	0.9	1159	4	US-09-181-585-1	Sequence 1, Appli
26	40.8	0.9	1471	4	US-09-181-585-2	Sequence 2, Appli
27	40.8	0.9	2095	4	US-09-620-312D-748	Sequence 748, App

QY	361	TTT	TAGCTCAAGATCATCTCAATTTATGTGCAAGTGTTGTAATTTCCCTCAAGACTATATAT	420
DB	361	TTT	TAGCTCAAGATCATCTCAATTTATGTGCAAGTGTTGTAATTTCCCTCAAGACTATATAT	420
QY	421	GAG	ATGTTTGTGTTTCATATTTCCAAATTTCAAACTTTGTCGCCATTTAGTCTTACCCCT	480
DB	421	GAG	ATGTTTGTGTTTCATATTTCCAAATTTCAAACTTTGTCGCCATTTAGTCTTACCCCT	480
QY	481	TCAT	CGATGGTTAGCTTTAGCTTTAATTCGTGAACCTGTTGAATTAACGATATGGCCCTTATGCT	540
DB	481	TCAT	CGATGGTTAGCTTTAGCTTTAATTCGTGAACCTGTTGAATTAACGATATGGCCCTTATGCT	540
QY	541	AAA	AGAAACAAACCTTATGGGTCATAAAAAAATAAGCCCAATATAAACTATAGGCCCAAA	600
DB	541	AAA	AGAAACAAACCTTATGGGTCATAAAAAAATAAGCCCAATATAAACTATAGGCCCAAA	600
QY	601	TAA	GTTTAGGTCATTTAGAGTGTGAGATAGCGCGTGTAGTGAACCGCACAGAAATGCGC	660
DB	601	TAA	GTTTAGGTCATTTAGAGTGTGAGATAGCGCGTGTAGTGAACCGCACAGAAATGCGC	660
QY	661	GTT	CGATTTGTTGGTGAAGTAGTCGTCTAGATTTCCCGGTCACCATGATGTTTCTTAGTGTAT	720
DB	661	GTT	CGATTTGTTGGTGAAGTAGTCGTCTAGATTTCCCGGTCACCATGATGTTTCTTAGTGTAT	720
QY	721	CAG	ACAGCTGTCGACAAACTGGTGGGAGAGATTAACGATCTTTAAGTAGGTCCCACTAGAT	780
DB	721	CAG	ACAGCTGTCGACAAACTGGTGGGAGAGATTAACGATCTTTAAGTAGGTCCCACTAGAT	780
QY	781	CAA	GATATTATAACCAATTGACCTTTTAACTTTTCAGGTAGTCCCGGAATCTGTGGCT	840
DB	781	CAA	GATATTATAACCAATTGACCTTTTAACTTTTCAGGTAGTCCCGGAATCTGTGGCT	840
QY	841	AGA	ATACAAAGAGGTTGTGAACAAAGTTGATGTTAAAGATGACACAAGAAATGTAATCTGAAC	900
DB	841	AGA	ATACAAAGAGGTTGTGAACAAAGTTGATGTTAAAGATGACACAAGAAATGTAATCTGAAC	900
QY	901	AAA	GCTGAATCATCTCTTCAGCCACTAGTATGTTTGACATATGCGCAGTTCCTTTGTAGC	960
DB	901	AAA	GCTGAATCATCTCTTCAGCCACTAGTATGTTTGACATATGCGCAGTTCCTTTGTAGC	960
QY	961	CTC	GAATAATAATAATAAAGTTTGAGGTTTAAAGATAATATAGTGGCTGAGATTCT	1020
DB	961	CTC	GAATAATAATAATAAAGTTTGAGGTTTAAAGATAATATAGTGGCTGAGATTCT	1020
QY	1021	CCA	TTCCGTAGCTCTCTGCTCTCTTTCTTTGTTTTCATTGATCAAAAGCAATCACTTCT	1080
DB	1021	CCA	TTCCGTAGCTCTCTGCTCTCTTTCTTTGTTTTCATTGATCAAAAGCAATCACTTCT	1080
QY	1081	TCT	CTCTCTCTCTCTCGAATTTCTACTGTTTCTTATCCACGAAATCTGGAATTAATAAA	1140
DB	1081	TCT	CTCTCTCTCTCTCGAATTTCTACTGTTTCTTATCCACGAAATCTGGAATTAATAAA	1140
QY	1141	TGG	AATCTTTATCGAATCCAAAGCTGATTTGTTCTTTTCATTGATCAATCTCTCTAAAG	1200
DB	1141	TGG	AATCTTTATCGAATCCAAAGCTGATTTGTTCTTTTCATTGATCAATCTCTCTAAAG	1200
QY	1201	TAC	TTAAGATTGATTTATGTCATGGTCTTTCTTATTTGTTTGAATGAATCACTGACTTGA	1260
DB	1201	TAC	TTAAGATTGATTTATGTCATGGTCTTTCTTATTTGTTTGAATGAATCACTGACTTGA	1260
QY	1261	TGT	TTTTTTGTTTGTGGAATTTAGTGGAAATTTTGTAAAGAGAAATCTGGAATTTGTPAGA	1320
DB	1261	TGT	TTTTTTGTTTGTGGAATTTAGTGGAAATTTTGTAAAGAGAAATCTGGAATTTGTPAGA	1320
QY	1321	GG	AGCTTAGTGATGAGACAAATTCGCTCGAGAGATCTGTTTATTTAAGGTAATTTAAC	1380
DB	1321	GG	AGCTTAGTGATGAGACAAATTCGCTCGAGAGATCTGTTTATTTAAGGTAATTTAAC	1380
QY	1381	TAA	ATTTTAGGGGAAGATGATGTTTGTAGGTGTCAAGATTTGAGAATTTTAATGAACT	1440
DB	1381	TAA	ATTTTAGGGGAAGATGATGTTTGTAGGTGTCAAGATTTGAGAATTTTAATGAACT	1440
QY	1441	TGA	TATAGACTCGGAAGCCATATACGATTAACAAACGAAACGTTGAAGGTGACCTGAGGAG	1500

[illegible]

Db 2521 AGNAGACAACTGTTTCAGATTGTTTCACTCATCAGTATCTCTGCTGCATCCTCCATGA 2580  
Qy 2581 ATAAAAGTTGTATAGACATCAAAACGCAAGCACTTCCCGAGTCTTCCGCTTTCACGGG 2640  
Db 2581 ATAAAAGTTGTATAGACATCAAAACGCAAGCACTTCCCGAGTCTTCCGCTTTCACGGG 2640  
Qy 2641 AAGAGTTAAAACAACTTCTTCATCTGCTATTTAGGTTTAAAGCAGTATAGTACTTTCA 2700  
Db 2641 AAGAGTTAAAACAACTTCTTCATCTGCTATTTAGGTTTAAAGCAGTATAGTACTTTCA 2700  
Qy 2701 TGAACCTAAAACCGTGGGGAATACAGGGAAGTCAGAATAACAGGGTAAAGAAAGGAGTC 2760  
Db 2701 TGAACCTAAAACCGTGGGGAATACAGGGAAGTCAGAATAACAGGGTAAAGAAAGGAGTC 2760  
Qy 2761 AAATCAGATTGAATGCAAAATCTCTGGAACCGGTAATGAGCAAGACCTCAGACTTA 2820  
Db 2761 AAATCAGATTGAATGCAAAATCTCTGGAACCGGTAATGAGCAAGACCTCAGACTTA 2820  
Qy 2821 TCCGATCATATCCCTGTGCTAGTGCATCCGCCACACAGTTGCGAGGATATCAGTCTGTTCC 2880  
Db 2821 TCCGATCATATCCCTGTGCTAGTGCATCCGCCACACAGTTGCGAGGATATCAGTCTGTTCC 2880  
Qy 2881 TCCTCTTCAGAGCCAGATAGTCAATCCGCCACACAGTTGCGAGGATATCAGTCTGTTCC 2940  
Db 2881 TCCTCTTCAGAGCCAGATAGTCAATCCGCCACACAGTTGCGAGGATATCAGTCTGTTCC 2940  
Qy 2941 TAATCATATAATGTCAACCTTTTACAAACACCGGCTCTTTATCTCCGCAACTTTTCG 3000  
Db 2941 TAATCATATAATGTCAACCTTTTACAAACACCGGCTCTTTATCTCCGCAACTTTTCG 3000  
Qy 3001 CTCAATATTTGGGCTCCCGATCTAGTGGTGGCTCAGCTGTTCCAGGGAACCTCACCTCC 3060  
Db 3001 CTCAATATTTGGGCTCCCGATCTAGTGGTGGCTCAGCTGTTCCAGGGAACCTCACCTCC 3060  
Qy 3061 GAATCTGGCTGCCATGGCCGAGCACTGTTGCACTGCTAGTCTGTTGGGGCTGCCAA 3120  
Db 3061 GAATCTGGCTGCCATGGCCGAGCACTGTTGCACTGCTAGTCTGTTGGGGCTGCCAA 3120  
Qy 3121 TGGATTATACCTTTATGTGCTCTCTTAGTTTCAAGTGGTTCATAGTCATCTCCATC 3180  
Db 3121 TGGATTATACCTTTATGTGCTCTCTTAGTTTCAAGTGGTTCATAGTCATCTCCATC 3180  
Qy 3181 TACTTTTGGACCATCATGTGATGTAGAGTACAAAAGCAAGCACTTTACAAATGTTTC 3240  
Db 3181 TACTTTTGGACCATCATGTGATGTAGAGTACAAAAGCAAGCACTTTACAAATGTTTC 3240  
Qy 3241 TGTGACAGCCGAGCAAGAACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTC 3300  
Db 3241 TGTGACAGCCGAGCAAGAACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTC 3300  
Qy 3301 AGAGGATGTTGAAATTAAGAGTAAACCAGTTTGTATGAGCAGCTTCTGCAACACCTGA 3360  
Db 3301 AGAGGATGTTGAAATTAAGAGTAAACCAGTTTGTATGAGCAGCTTCTGCAACACCTGA 3360  
Qy 3361 GAGTGATGCAAGGGTTTCAGATGAGCAGGAGCAGAAACAAAGTTACCGGCTCTCGTG 3420  
Db 3361 GAGTGATGCAAGGGTTTCAGATGAGCAGGAGCAGAAACAAAGTTACCGGCTCTCGTG 3420  
Qy 3421 TGGCTCAAACTCCGTCGAGTGTATGATGTTGAGCGGATGATGATGTTGAGCGGATGATG 3480  
Db 3421 TGGCTCAAACTCCGTCGAGTGTATGATGTTGAGCGGATGATGATGTTGAGCGGATGATG 3480  
Qy 3481 GGATGGCAACATGTTGAGGTGAAAGAAACGAATGAAGACACTAATAAACCCTCAAACTTC 3540  
Db 3481 GGATGGCAACATGTTGAGGTGAAAGAAACGAATGAAGACACTAATAAACCCTCAAACTTC 3540  
Qy 3541 AGAGTCCAATGCAAGCGGAGTGAATCAGCTCAATATACCGATCCATGGAAGTCTGT 3600  
Db 3541 AGAGTCCAATGCAAGCGGAGTGAATCAGCTCAATATACCGATCCATGGAAGTCTGT 3600  
Qy 3601 GTCTGACGAGTACTTACTTGGACTAAAGATCAACTTCTTTTATTTCAAAATCAATTTTCTC 3660  
Db 3601 GTCTGACGAGTACTTACTTGGACTAAAGATCAACTTCTTTTATTTCAAAATCAATTTTCTC 3660

Query Match 35.2% Score 1527.6; DB 4; Length 2254;

## RESULT 2

US-08-843-572E-3

; Sequence 3, Application US/08843572E

; Patent No. 6388172

; GENERAL INFORMATION:

; APPLICANT: TOBIN, ELAINE

; APPLICANT: WANG, ZHI-YONG

; APPLICANT: SUN, LIN

; TITLE OF INVENTION: PHYTOCHROME REGULATED

; TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR CONTROL OF HIGHER PLANT

; FILE REFERENCE: p601-D-36315.60100

; CURRENT APPLICATION NUMBER: US/08/843,572E

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: Microsoft Word 7.0 (DOS text file)

; SEQ ID NO 3

; LENGTH: 2254

; TYPE: DNA

; ORGANISM: Arabadopsis/thalia

US-08-843-572E-3

[illegible]



```

; Sequence 6, Application US/09202712
; Patent No. 6265637
; GENERAL INFORMATION:
; APPLICANT: Coupland, George M
; APPLICANT: Schaffer, Robert J
; TITLE OF INVENTION: Genetic control of flowering
; FILE REFERENCE: 620-55
; CURRENT APPLICATION NUMBER: US/09/202,712
; CURRENT FILING DATE: 1999-01-11
; EARLIER APPLICATION NUMBER: PCT/GB97/01676
; EARLIER FILING DATE: 1997-06-23
; EARLIER APPLICATION NUMBER: GB 9613132.1
; EARLIER FILING DATE: 1996-06-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant
US-09-202-712-6

Query Match
Best Local Similarity 2.1%; Score 90.4; DB 3; Length 666;
Matches 193; Conservative 1; Mismatches 77; Indels 37; Gaps 2;

QY 1444 TATAGACTCGGAAGCCATATACGATAACAAAGCAACGTGAAAGGTGGACTGAGGAAGAAC 1503
DB 284 TTATGGCAAGAAAGCCATATACATACAAACAGCAGGAGCGATGGACTGAGGATGACC 343
QY 1504 ATATAGATTTCATTGAGCTTTGAGGCTTTATGGTAGAGCATGGGAGAGATTTGAAGTT 1563
DB 344 ATGAGAGGTTCTAGAGGCTTTGAGGCTTTATGGAAGAGCTTTGGCAACGAATTGAAGTTC 403
QY 1564 G-----ATTTTATTTCCCTTATATCTCTATTTTGTGTTGAG 1606
DB 404 GRAAGTTTATCTTTTGAATGTTAGTTGAACTCTTTGAGATTTTATATCTCTTTGTTT 463
QY 1607 AGGTTTGTCTTCA-----AACTGATTTGCTTTTTCATTGGAC 1646
DB 464 AGGAGTGCTTTATCTCTCTGTTGAGGAGATTCCTTTCTTTTCATTGTTGTC 523
QY 1647 AGAATATAGTACCAAAACAACTGCTGTCAGATAGAGTACGCTCAGAAATTTTCTC 1706
DB 524 AGAATATATGGGACAAAGACTGCTGTTGAGATCAGAAATCATGCAAAAGTTCTTCAC 583
QY 1707 CAAGGTAA 1714
DB 584 AAAGGTAA 591

RESULT 4
US-09-202-712-1
; Sequence 1, Application US/09202712
; Patent No. 6265637
; GENERAL INFORMATION:
; APPLICANT: Coupland, George M
; APPLICANT: Schaffer, Robert J
; TITLE OF INVENTION: Genetic control of flowering
; FILE REFERENCE: 620-55
; CURRENT APPLICATION NUMBER: US/09/202,712
; CURRENT FILING DATE: 1999-01-11
; EARLIER APPLICATION NUMBER: PCT/GB97/01676
; EARLIER FILING DATE: 1997-06-23
; EARLIER APPLICATION NUMBER: GB 9613132.1
; EARLIER FILING DATE: 1996-06-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2526
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)..(2275)
US-09-202-712-1

Query Match
Best Local Similarity 1.9%; Score 83.2; DB 3; Length 2526;
Matches 233; Conservative 0; Mismatches 213; Indels 12; Gaps 1;

QY 2847 CATTGGGAGCTCAATACAAAGTTCTTATCATCATCTCTCTCAGAGCCAGATAGTATC 2906
DB 1137 CATCTGCTACTACTACAGCTTCTCATCAAGCGCTTTCCAGCTTGTCAATCACAGGATG 1196
QY 2907 CCCACACAGTTGCAGGAGATTATAGTCGTTTCTTAATCATATAATGTCAACCTTTTAC 2966
DB 1197 ACCGTTTCGTTCTCCAGATATCATCTACTTTCTCAATCTTATATGTCAACTCTCTAC 1256
QY 2967 AAACACCGGCTCTTTATATCTGCGCAACTTTCCGCTCATCATTTTGGCCCTCCCG----- 3020
DB 1257 AGAATCCTGCAGCTCATGCTGCAGCTACATTCGCTGCTTCGGCTTATGCGAGTG 1316
QY 3021 -----ATTCTAGTGGTGGCTCACCTGTTCCAGGAACTCACCTCGAATCTGGCTGCCA 3074
DB 1317 TCGGAAATTCGGTGATTCATCAACCCCAATGAGCTCTTCTCTCAAGTATAAATGCGCA 1376
QY 3075 TGGCCGAGCCACTGTTGCAGCTGCTAGTGTGTTGGGCTGCCAATGATTAATACCTT 3134
DB 1377 TTGCGGCTGTACAGTAGCTGCTGCACTGCTTTGGTGGCTTCTCATGACTTCTTCTCTG 1436
QY 3135 TATGTGCTCTCTTATGTTAGTTCAGTGTTTCACTAGTATCTCTCCATCTACTTTTGGACCAT 3194
DB 1437 TATGCGCTCCAGCTCCAATAACATGTGTTCCATTTCTCAACTGTGTCAGTTTCCAACCTCCAG 1496
QY 3195 CATGTGATGTAGAGTACACAAAGCAAGCACTTTACACATCGTTCTGTGCGAGCCGAG 3254
DB 1497 CAATGACTGAAATGATACCGTTGAAATACTCAACCGTTTGAGAAACAAACACAGCTC 1556
QY 3255 AGCAAGAACACTCCAGGAGCATCAAAAGGCTCGATCTTCA 3292
DB 1557 TGCAGATCAAACTTGGCTTCAATCTCCAGCTTCA 1594

RESULT 5
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991

```



```

; SOFTWARE: PatentIn Release #1.0, Version #1.2.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/487,826B
;
; FILING DATE: 10-SEP-1993
;
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Israelsen, Ned
;
; REGISTRATION NUMBER: 29,655
;
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (619) 235-8550
;
; TELEFAX: (619) 235-0176
;
; INFORMATION FOR SEQ ID NO: 13:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 19124 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: cDNA
;
; HYPOTHETICAL: NO
;
; ANTI-SENSE: NO
;
; US-08-487-826B-13

```

Query Match	1.2%	Score 51.2;	DB 2;	Length 19124;
Best Local Similarity	46.7%;	Pred. No. 0.012;		
Matches 198;	Conservative 0;	Mismatches 223;	Indels 3;	Gaps 1;
QY	965	AAATAAATAAAATPAAAAAGTTTTGAGCTTAAAGATAATTATAGCTGGCTGAGATTTCTCCAT	1024	
Db	16004	ATATATATATATATATGATGATATATGCGCATGTAATTAATTTTATTTATTTTATTTTAT	15945	
QY	1025	TTCCGTAGCTTCGGTCTCTTTTCTTTGTTTCATTGATCAAAAGCAAAATCACTTCTCTT	1084	
Db	15944	TTAATTAATAATTTTTTATTTATTTTATTTTCATTAATAATTTTTTAAATTTTTTATTTA	15885	
QY	1085	CTTCTTCTCTCGATTCCTACTGTTTCTCTATCCAAGCAAAATCGGAATTAATAAATGGA	1144	
Db	15884	ATTAAATTTTTTATTTATTTATTTTATTTTATTTTATTTTAAATTTTATTTATTTATTTT	15825	
QY	1145	ATCTTTATCGAATCCAAGCTGATTTTGTTTCTTTCATTGAATCATCTCTCTAAAGGTACT	1204	
Db	15824	TTTTTATTTTAAAT--AAATTTTTTTTATTTTATGATATATTTTTTTTTTAAACATTTT	15768	
QY	1205	TAAGATTGATTTATTGTCATGCTCTTCTTATTTGTTGATGNAATCACTGACTTGATGT	1264	
Db	15767	TTAATTTTTTTTATTTTATGATATAATTTTTTATTTTAAATATATTTTTTCTTTTTTTT	15708	
QY	1265	TTTTTGTTTTTGGAATTAGTGGAAATTTTGTATAAGAGAAAGATCTGAAGTTGTAGAGGAG	1324	
Db	15707	TGTGTTTTATGATATATATTTTTTTTTTTTTTTTTTAAATGTTTTTTTTTCTCTTTGTTTTT	15648	
QY	1325	CTTAGTGATGGACAAAAATTCGCTCGGAGAGATCTGTTATTAGGTAAATTAACATAA	1384	
Db	15647	ATTTTTTTTATATCAATTTTTTTTTTTATATAAAAAATTTTTTTTAAATTTTTTTTGATAAT	15588	
QY	1385	TTTTT 1388		
Db	15587	TTTT 15584		

```

RESULT 8
US-09-202-712-19
; Sequence 19, Application US/09202712
; Patent No. 6265637
; GENERAL INFORMATION:
; APPLICANT: Coupland, George M
; APPLICANT: Schaffer, Robert J
; TITLE OF INVENTION: Genetic control of flo
; FILE REFERENCE: 620-55
; CURRENT APPLICATION NUMBER: US/09/202,712
; CURRENT FILING DATE: 1999-01-11
; EARLIER APPLICATION NUMBER: PCT/GB97/01676
; EARLIER FILING DATE: 1997-06-23

```

```

; EARLIER APPLICATION NUMBER: GB 9613132.1
; EARLIER FILING DATE: 1996-06-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Nucleotide
; OTHER INFORMATION: sequence of EST 162131T7, Accession no. R30439
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (37, 204, 275, 335, 347, 353, 363, 367)
; OTHER INFORMATION: n is a, g, c, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (377, 384, 392, 393, 419, 428, 429, 436, 439)
; OTHER INFORMATION: n is a, g, c, or t
US-09-202-712-19

```

	Query Match	1.1%	Score 46.2;	DB 3;	Length 447;
	Best Local Similarity	62.1%;	Pred. No. 0.04;		
	Matches	72;	Conservative	0;	Mismatches 44; Indels 0; Gaps 0;
Qy	1453	GGAAGCCNATACGATTACAAGCAACGTGAAAGGTGGACTGAGGAGAAGAACATAATAGAT			1512
Db	150	GGAAGCTTACACAATCACCACCTAGAGAGAGTTGGACTGAAGGAGAACCGNCAAGT			209
Qy	1513	TCATTGAAGCTTTGAGGCCTTTATGGTAGACATGGCAGAGAAGATTGAAGGTGATTT			1568
Db	.210	TTCCTGAAGCTCTCAAATGTGTTTGATCGTGACTGGAAAAGAGATAGAAGATTTTTTT			265

## RESULT 9

```

US-09-523-899A-1/c
; Sequence 1, Application US/09523899A
; Patent No. 6544789
; GENERAL INFORMATION:
; APPLICANT: Raskin, Ilva
; APPLICANT: Haran, Shoshan
; TITLE OF INVENTION: Phosphorus-Cont
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 29155/35950
; CURRENT APPLICATION NUMBER: US/09/52
; CURRENT FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 09/015,601
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1274
; TYPE: DNA
; ORGANISM: A. thaliana
; FEATURE:
; OTHER INFORMATION: A. thaliana acci
US-09-523-899A-1

```

	Query Match Best Local Similarity Matches 243; Conservative	1.0%; 45.3%;	Score 45.6; Pred. No. 0.09;	DB 4; Mismatches 289;	Length 1274; Indels 4; Gaps 2;
QY	980	AAAGTTTGAGGTTAAAGATAATTATAGTCGGCTGAGATTCTCCATTTCGGTAGCTCTCTGG	1039		
Db	851	AAAGTTTCTCGAAGAGAAACATCTCTCCTTCTCTCATATTTTCTCTCTCTCTCA	792		
QY	1040	TCCTTTCTCTTTGTTTCATGTATCAAAAGCAAAATCACTTCTCTCTCTCTCTCTCGAT	1099		
Db	791	TATTTTTTTTTTTGGGTATCTCTCATTTGAGATCTCCAAATGAAATGCTCTAACTATCT	732		
QY	1100	TTCTTACTGTTTTCTTATCCACGAAATCTCGAAATTAAAAATGGAATCTTTATCGAATCC	1159		
Db	731	TTTTTTTTTTTTTAAATATCGATTAAAAAGTGTTGTAGGAAA--CAATTGATAGTGTGCT	674		

Qy	1160	AAGCTGATTTTGTCTTTTCATTGCAATCATCTCTCTAAAGGTACTTAAAGATTGATTTATT	1219
Db	673	TTCATTTTTTCTTTACAGATTTTAAAAACAAGATTTTGAATGTA--AAATAAGCTTTTTTTT	616
Qy	1220	GTCATGCTTTTCTTATTTGTTTGATGAATAACTTGACTCGATTGTTTTTTTTTTGTTTGGA	1279
Db	615	AAAAATATATTTTCATTACCTTTTGTTTAGGAAGGTTATTGGTTTATAGATTTATAAT	556
Qy	1280	TTAGTGGAAATTTGTAAAGAGAAGACTGAAGTTTGTCTAGAGGAGCTTACTGATGGAGAC	1339
Db	555	GAATTGAAATCCGAACAAAAATCATTTGTAATTTAATTTGATTTTAAAAATTTTAAT	496
Qy	1340	AAATTCTCTGGAGAAGACTGTTTATTAAAGTTAAATTACTAAATTTTAGGGGGGAAGAT	1399
Db	495	AATATCATGTGTTATTCATTCAATGATTTTATAAATCTAGTTAAAAATCCCTTTGTTATTCG	436
Qy	1400	GATTGTTTTTAGTGCTCAAAGATTGAGAAATTTTAAATGAAACTTGATATAGACTCGGAAGCC	1459
Db	435	TTTAGTCATTGTTTTTTGGATTGAGAAATCCAAATTAATGTTTTAAATAAAATTAATGGA	376
Qy	1460	ATATACGATTACAAAGCAACGTGAAAGGTGGACTGAGGAAGAACATAATAGATTCA	1515
Db	375	TTTGAAGAGAAAAATAGAAGATTTCAAATCCATGGGATAAACACATATTTATTAA	320

RESULT 10

```

US-08-623-906A-18/c
; Sequence 18, Application US/08623906A
; Patent No. 5874217
; GENERAL INFORMATION:
; APPLICANT: Stevenson, Tamara
; APPLICANT: Dvorak, Jan
; APPLICANT: Halverson, Joy
; TITLE OF INVENTION: Microsatellite Sequences for Canine
; TITLE OF INVENTION: Genotyping
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/623,906A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A-62282/BIIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..79
; OTHER INFORMATION: /note= "Nucleotides 1-79 are unique"
; OTHER INFORMATION: flanking sequence"
; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: 80..229
; OTHER INFORMATION: /note= "Nucleotides 80-229 are
; OTHER INFORMATION: repeat sequence"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 230..376
; OTHER INFORMATION: /note= "Nucleotides 230-376 are
; OTHER INFORMATION: unique flanking sequence"
US-08-6233-906A-18

Query Match      1.0%; Score 45.4; DB 2; Length 376;
Best Local Similarity 47.1%; Pred. No. 0.06;
Matches 133; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 1013 AGATTTTCTCCATTCCGTAGCTTCTGGTCTCTTTTCTTTGTTTCATTGATCAAAAGCAA 1072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 292 AGAAGCTTCAGATTATTTATTTCTATTTCGTGTCTAAGTCATTAAATCTCTCCTCGA 233

Qy 1073 TCACCTCTCTCTCTCTCTCTCTCTCGAATTCCTTACTGTTCCTTATCAACGAAATCTGGA 1132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 ATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTCTC 173

Qy 1133 ATTAAGAAATCGAATCTTTATCGAATCCAGCTGATTTTGTTTCTTCATTGAATCATCTC 1192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 113

Qy 1193 TCTAAGAGTACTTAAAGATTGATTTATTTGTGCATGGTCTTTTCTTATTTGTTGATCAATAACT 1252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 TCITTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 53

Qy 1253 TGACTTGATTGTTTTTGTGTTTGTGGAATGATGGAATTT 1291
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 52 TATTCCAAATATTCCAAATTTTCGAAATTTTCGAAATTCGTTTATCTT 14

```

RESULT 11

US-07-867-106-2  
Sequence 2, Application US/07867106  
Patent No. 5389526  
GENERAL INFORMATION:  
APPLICANT: Slade, Martin B  
APPLICANT: Chang, Andy C M  
APPLICANT: Williams, Keith L  
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris  
STREET: One Liberty Place 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/867,106  
FILING DATE: 19920625  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PJ 7187  
APPLICATION NUMBER: PCT/AU90/00530  
FILING DATE: 02-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Feeney, Joanne Longo  
REGISTRATION NUMBER: 35,134  
REFERENCE/DOCKET NUMBER: RICE-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439



Query Match 1.0%; Score 42.2; DB 1; Length 4380;  
Best Local Similarity 50.2%; Pred. No. 1.1;  
Matches 130; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

1031	AGCTTCGGTCTCTTTCTTTGTTTTCATTCATCAAAAGCA - AATCACTTCTCTTCTCTTCT	1089
y		
b		
344	AAC TTC TCTCCATTTTACATGTTTAAATGCTGT CAGGATGATCAGTACACAACTTCTCT	285
1090	TCTTCTCGATTCTTACTGTTTCTTATCCAAAGAAATCTGGAATTAATAATGGAATCTT	1149
y		
b		
284	AAAAATTTTCATCGGCACCTTTTTCCTTTTACAGTAAATCTGAGTTAAAAAATGTTTGT	225
1150	TATCGAATCCAAAGCTGATTTGTTTCTTTTCATTGAATCATCTCTTAAAGGTACTTTAAGA	1209
y		
b		
224	TTTCATATCCATAACCTCTATGTTAGTTTATTCCTAAAGTAAGATAATGCTTACTTAAATA	165
1210	TTGATTATTTCATGGTCTTCTTCTTATTGTTGATGAATAAATCTGATCTTCTTCTTCTT	1269
y		
b		
164	AACATATTAAATGTTATTTGGTTATAAATAAAACCCCAACATTCGCAGTGTTTGGGTTTA	105
1270	GTTTTTGGATTAGTGAA	1288
y		
104	TTTTTTTGATGTAATAAAA	86

Search completed: November 24, 2003, 13:07:50  
Job time : 239.433 secs

RESULT IS  
US-07-582-945-1/C

GenCore version 5.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 18:53:22 ; Search time 8051.27 Seconds  
 (without alignments)  
 11452.890 Million cell updates/sec

Title: US-10-084-553-3

Perfect score: 2254

Sequence: 1 t9agattctctccatttcctgt.....tttccaaaaaaaaaaaaaa 2254

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb\_ba.\*  
 2: gb\_hgc.\*  
 3: gb\_in.\*  
 4: gb\_om.\*  
 5: gb\_ov.\*  
 6: gb\_pat.\*  
 7: gb\_ph.\*  
 8: gb\_pl.\*  
 9: gb\_pr.\*  
 10: gb\_ro.\*  
 11: gb\_sts.\*  
 12: gb\_sv.\*  
 13: gb\_un.\*  
 14: gb\_vi.\*  
 15: em\_ba.\*  
 16: em\_fun.\*  
 17: em\_hum.\*  
 18: em\_in.\*  
 19: em\_mu.\*  
 20: em\_om.\*  
 21: em\_or.\*  
 22: em\_ov.\*  
 23: em\_pat.\*  
 24: em\_ph.\*  
 25: em\_pl.\*  
 26: em\_ro.\*  
 27: em\_sts.\*  
 28: em\_un.\*  
 29: em\_vi.\*  
 30: em\_hgc\_hum.\*  
 31: em\_hgc\_inv.\*  
 32: em\_hgc\_other.\*  
 33: em\_hgc\_mus.\*  
 34: em\_hgc\_pln.\*  
 35: em\_hgc\_rod.\*  
 36: em\_hgc\_nam.\*  
 37: em\_hgc\_vrt.\*  
 38: em\_sy.\*  
 39: em\_hgo\_hum.\*  
 40: em\_hgo\_mus.\*  
 41: em\_hgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2254	100.0	2254	6	BD107661 Phytochro
2	2240	99.4	2240	6	U28422 Arabidopsis
3	1827	81.1	1827	6	AX506805 Sequence
4	1805	80.1	2684	8	AY050961 Arabidops
5	1612	71.5	1612	8	BT001096 Arabidops
6	1612	71.5	1612	8	BT001105 Arabidops
7	1527.6	67.8	4344	6	BD107660 Phytochro
8	1527.6	67.8	5242	6	U7156 Arabidopsis
9	1526.6	67.7	68415	8	AC005310 Arabidops
10	223.4	9.9	318	8	AY166504 Brassica
11	217.2	9.6	2526	8	A68428 Sequence 1
12	217.2	9.6	2526	8	ATAJ6404 Arabidops
13	213	9.4	1938	6	AX506421 Sequence
14	202.2	9.0	2172	8	PVU420902 Phaseolus
15	189	8.4	1134	6	AX509497 Sequence
16	137.6	6.1	1072	8	AY114070 Arabidops
17	137.6	6.1	1429	8	AY063952 Arabidops
18	117.6	5.2	2160	6	AX660076 Sequence
19	112	5.0	451	6	AX660581 Sequence
20	109.8	4.9	1317	6	AX658868 Sequence
21	105.4	4.7	1638	6	AX658770 Sequence
22	104.2	4.6	1332	8	AY087690 Arabidops
23	104.2	4.6	1467	8	BT002002 Arabidops
24	97.8	4.3	928	8	AY122969 Arabidops
25	97.8	4.3	1427	8	AY091106 Arabidops
26	94.8	4.2	2196	6	AX658772 Sequence
27	90.4	4.0	999	6	AX658842 Sequence
28	90	4.0	510	6	AX658822 Sequence
29	81.6	3.6	86436	8	AC007323 Genomic s
30	81.4	3.6	1051	8	AY081522 Arabidops
31	81.4	3.6	1335	8	AY042894 Arabidops
32	79.4	3.5	447	6	A68446 Sequence 19
33	75.4	3.3	666	6	A68433 Sequence 6
34	70	3.1	14826	8	AP004460 Oryza sat
35	69	3.1	124212	8	AC126787 Medicago
36	67.6	3.0	131006	2	AC137985 Medicago
37	65.8	2.9	139490	8	AP003225 Oryza sat
38	64.6	2.9	95327	8	AC013354 Genomic s
39	64.6	2.9	96232	8	ATAC010927 Arabidops
40	63	2.8	186315	8	OSJN00188 Arabidops
41	63	2.8	229796	2	AC137597 Oryza sat
42	62.8	2.8	151517	2	AP005750 Oryza sat
43	62.8	2.8	167114	2	AP004685 Oryza sat
44	59.8	2.7	75188	8	AB005238 Arabidops
45	59.2	2.6	7218	6	I66494 Sequence 14

## ALIGNMENTS

RESULT 1  
 BD107661  
 LOCUS  
 DEFINITION  
 Phytochrome regulated transcription factor for control of higher plant development.  
 ACCESSION  
 BD107661  
 VERSION  
 BD107661.1 GI:23202479  
 KEYWORDS  
 JP 2002501381-A/2.  
 SOURCE  
 unidentified.  
 ORGANISM  
 unclassified.  
 REFERENCE  
 1 (bases 1 to 2254)  
 Tobin, R., Sun, L. and Wang, Z.Y.  
 TITLE  
 Phytochrome regulated transcription factor for control of higher plant development

BD107661 2254 bp DNA linear PAT 18-SEP-2002  
 Phytochrome regulated transcription factor for control of higher plant development.





```

QY 1741 TTGCGCAAGTTTTACATATCGAGAGAACACACAGAGAGGAGAACACAAACAAGAA 1800
Db 1741 TTGCGCAAGTTTTACATATCGAGAGAACACACAGAGAGGAGAACACAAACAAGAA 1800
QY 1801 CAAAGATATCCAATGGACCTGATCTTTAACTTTCACAGCTCAGTTAAACACAGTTGATGAT 1860
Db 1801 CAAAGATATCCAATGGACCTGATCTTTAACTTTCACAGCTCAGTTAAACACAGTTGATGAT 1860
QY 1861 CAAGAGAGAGAGAGAAACACAGGATTTCTTGGATCGGATTTAGATGCTTCAAGCTAATG 1920
Db 1861 CAAGAGAGAGAGAGAAACACAGGATTTCTTGGATCGGATTTAGATGCTTCAAGCTAATG 1920
QY 1921 AGTAGAGAGAGAGAGAGGTTTTTAACCATACAAAGATGTTCCATGGAGGCAAGAAAGT 1980
Db 1921 AGTAGAGAGAGAGAGAGGTTTTTAACCATACAAAGATGTTCCATGGAGGCAAGAAAGT 1980
QY 1981 AGAATCTCTCAACCAACATCTCTATCATTCATGTCGGAACAGAAAGATCCCAACGGATCGG 2040
Db 1981 AGAATCTCTCAACCAACATCTCTATCATTCATGTCGGAACAGAAAGATCCCAACGGATCGG 2040
QY 2041 TTGGAACCTCAAGCTTCCACATGAGACTCTATTTTCATCTGATCTGTTTGTACTCTG 2100
Db 2041 TTGGAACCTCAAGCTTCCACATGAGACTCTATTTTCATCTGATCTGTTTGTACTCTG 2100
QY 2101 TTTTAAAGTTTTCAAGACCACTGCTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2160
Db 2101 TTTTAAAGTTTTCAAGACCACTGCTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2160
QY 2161 TTCTTGTCTCATTAGTCTTCTCTGTAACATTTGACTCTGATTTATTTCAACAATCATAACT 2220
Db 2161 TTCTTGTCTCATTAGTCTTCTCTGTAACATTTGACTCTGATTTATTTCAACAATCATAACT 2220
QY 2221 GTTTAAATCTTTTTTTTCCAAAAAATAAAAAA 2254
Db 2221 GTTTAAATCTTTTTTTTCCAAAAAATAAAAAA 2254

```

```

RESULT 2
ATU28422 2240 bp mRNA linear PLN 28-SEP-2001
LOCUS Arabidopsis thaliana DNA-binding protein CCA1 (CCA1) mRNA, complete
DEFINITION cds.
ACCESSION U28422
VERSION U28422.1 GI:1777442
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

```

```

REFERENCE
AUTHORS Wang,Z.Y., Kenigsbuch,D., Sun,L., Harel,E., Ong,M.S. and Tobin,E.M.
TITLE A Myb-related transcription factor is involved in the phytochrome
regulation of an Arabidopsis lhcb gene
JOURNAL Plant Cell 9 (4), 491-507 (1997)
MEDLINE 97290202
PUBMED 914958

```

```

REFERENCE
AUTHORS Wang,Z.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-1995) Zhiyong Wang, Dept. of Biology, University
of California at Los Angeles, 405 Hilgard Ave., Los Angeles, CA
90095, USA

```

```

FEATURES
Source
1..2240
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/dev_stage="21-day-old green seedlings"
1..2240
/gene="CCA1"
238..2064
CDS

```

```

/gene="CCA1"
/function="DNA binding protein"
/note="myb-related DNA binding protein"
/codon_start=1
/product="CCA1"
/protein_id="AAB40525.1"
/db_xref="GI:1777443"
/translation="METNSGDELIVIKTRPYTIKORWTEBEHRNFIABRLRYPGR
AWOKIEBHVAKTAVQIRSHAKOFFSKVEAEAKGVAMGOALDIAIPPPRKRKPNR
PYPRKTSGTILMSKTCVNDGKSLGSEKVSHPMANEDROQSPBEKTLQEDNCSDC
FTQYLSAASSMKNKSIETSNATREFRFSREGRSQNNRVKESNDLNKSLNENGN
EQQPTVPMHIVPLVPLUGSSITSLSHPPSEDPSPHPTVAGDYOSFNHINSTLLQTP
ALYTAATFASFPFPPDSSGSPVGNPNLAAAMAAATAVAASAAWANGLLPLCAPL
SSGGFTSHPPSTFTGSCDVEYTKASTLQHGVSQREHSEASKARSLDSEDEVKLS
KPYCHEQPSATPESDAGKDRKQVDRSSCGSNTPSSDDVDEADAEERQEDQING
EVETNEDTNKPTSESNAARRISNIITDPMKSVSDGRIAFQALFSEVLPVPSFTY
REHREEEQQQEQRYPMALDLNFTAQLTTPVDDQEKRNFTGLGIGLDASKUMSRGT
GFAPYKRCSEAKESRILNNPIIHVEQKQPKRMRLTQAST"
polyA_site
/gene="CCA1"
/note="14 A nucleotides"
BASE COUNT 685 a 473 c 477 g 605 t
ORIGIN

```

```

Query Match 99.4%; Score 2240; DB 8; Length 2240;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGAGATTCTCCATTCCTCGTAGCTTCCTGCTCTCTTTCTTTGTTTCATTGATCAAAAGCA 60
Db 1 TGAGATTCTCCATTCCTCGTAGCTTCCTGCTCTCTTTCTTTGTTTCATTGATCAAAAGCA 60
QY 61 AATCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 61 AATCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 121 GAATTAATAATGGAATCTTTATCGAATCCAAAGCTGATTTGTTTCTTTTCATTGATCATC 180
Db 121 GAATTAATAATGGAATCTTTATCGAATCCAAAGCTGATTTGTTTCTTTTCATTGATCATC 180
QY 181 TCTCTAAAGTGGAAATTTTGTAAAGAGAGATCTGAAGTTGTGTAGAGAGAGCTTGTAGTATG 240
Db 181 TCTCTAAAGTGGAAATTTTGTAAAGAGAGATCTGAAGTTGTGTAGAGAGAGCTTGTAGTATG 240
QY 241 GAGACAAATTCGCTCTGAGAGAGATCTGGTTATTAAAGACTCGGA;CCATATACCATACA 300
Db 241 GAGACAAATTCGCTCTGAGAGAGATCTGGTTATTAAAGACTCGGAAGCATATACCATACA 300
QY 301 AAGCAACGTGAAAGGTGGACTGAGGAAGAACATAATAGATTTCATTGAAGCTTTTGAGGCTT 360
Db 301 AAGCAACGTGAAAGGTGGACTGAGGAAGAACATAATAGATTTCATTGAAGCTTTTGAGGCTT 360
QY 361 TATGTTAGAGCATGGCAGAGAGATTGAAGAACATGTAGCAACAAAGCTGTGTCCAGATA 420
Db 361 TATGTTAGAGCATGGCAGAGAGATTGAAGAACATGTAGCAACAAAGCTGTGTCCAGATA 420
QY 421 AGAAGTCACGCTCAGAAAATTTTTCTCCAAGTAGAGAAAGAGGCTGAAGCTTAAGGTGTA 480
Db 421 AGAAGTCACGCTCAGAAAATTTTTCTCCAAGTAGAGAAAGAGGCTGAAGCTTAAGGTGTA 480
QY 481 GCTATGGGTCAAGCGCTAGACATAGCTATTTCTCTCTCCAGCGCTTAAGCGCTAAACCAAC 540
Db 481 GCTATGGGTCAAGCGCTAGACATAGCTATTTCTCTCTCCAGCGCTTAAGCGCTAAACCAAC 540
QY 541 AATCCTTATCTCGAAAGACGGGAAGTGGAAACGATCTTTATGTCAAAAACGGGTGTAAT 600
Db 541 AATCCTTATCTCGAAAGACGGGAAGTGGAAACGATCTTTATGTCAAAAACGGGTGTAAT 600
QY 601 GATGGAAGAGAGTCCCTTCGATCAGAAAAGTGTCTCCATCTCTGAGATGGCCATGAGAT 660
Db 601 GATGGAAGAGAGTCCCTTCGATCAGAAAAGTGTCTCCATCTCTGAGATGGCCATGAGAT 660
QY 661 CGACAAATCAAAAGCCTCGAAGAGAAACTCTCGAGGAGAACTCTGAGATTGTTTTCAGATTGTTTC 720

```

```
Db 661 CGACAAATCAAGCCTGAAGAAACTCTGCAGGAAGCAACTGTTTCAGATTGTTTC 720
Qy 721 ACTCATCAGTATCTCTCTGCTGATCTCTCATGATTAAGTTGTATAGAGACATCAAAC 780
Db 721 ACTCATCAGTATCTCTCTGCTGATCTCTCATGATTAAGTTGTATAGAGACATCAAAC 780
Qy 781 GCAAGCACTTTCCGCGAGTTCTTGGCTTTCACGGGAAGAGGGAAGTCAAGTAACAGGGTA 840
Db 781 GCAAGCACTTTCCGCGAGTTCTTGGCTTTCACGGGAAGAGGGAAGTCAAGTAACAGGGTA 840
Qy 841 AGAAGAGGAGTCAAACTCAGATTTGAATGCAAAATCTCTGGAACCGTAAATAGCAAGGA 900
Db 841 AGAAGAGGAGTCAAACTCAGATTTGAATGCAAAATCTCTGGAACCGTAAATAGCAAGGA 900
Qy 901 CCTCAGACTTATCCGATGCAATATCCCTGTGCTAGTGCCATTTGGGAGCTCAATAACAAGT 960
Db 901 CCTCAGACTTATCCGATGCAATATCCCTGTGCTAGTGCCATTTGGGAGCTCAATAACAAGT 960
Qy 961 TCTCTATCAGATCTCTCTCAGAGCCAGATAGTATCCCAACAGTTCGAGGAGATTAT 1020
Db 961 TCTCTATCAGATCTCTCTCAGAGCCAGATAGTATCCCAACAGTTCGAGGAGATTAT 1020
Qy 1021 CAGTCGTTCTTAATCATATAATGTCAACCTTTTACAAACACCGGCTCTTTTACTGCC 1080
Db 1021 CAGTCGTTCTTAATCATATAATGTCAACCTTTTACAAACACCGGCTCTTTTACTGCC 1080
Qy 1081 GCAACTTTCCGCTCATCATTTTGGCTTCCCGATTTCTAGTGGCTCACCTGTCAGGG 1140
Db 1081 GCAACTTTCCGCTCATCATTTTGGCTTCCCGATTTCTAGTGGCTCACCTGTCAGGG 1140
Qy 1141 AACTCACTCCGAATCTGGCTGCAATGGCCGCGACCACTGTTGAGCTGCTAGTGTGG 1200
Db 1141 AACTCACTCCGAATCTGGCTGCAATGGCCGCGACCACTGTTGAGCTGCTAGTGTGG 1200
Qy 1201 TGGGCTGCCAATGATTAATACCTTTATGTGCTCTCTTAGTTAGTGTGTTTCACTAGT 1260
Db 1201 TGGGCTGCCAATGATTAATACCTTTATGTGCTCTCTTAGTTAGTGTGTTTCACTAGT 1260
Qy 1261 CATCTCCATCTACTTTTGGACCATCATGTGATGTAGTACACAAAGCAAGCACTTTA 1320
Db 1261 CATCTCCATCTACTTTTGGACCATCATGTGATGTAGTACACAAAGCAAGCACTTTA 1320
Qy 1321 CAACATGGTTCTGTGACAGCCGAGAGCAAGAACACTCCGAGGCATCAAGGCTCGATCT 1380
Db 1321 CAACATGGTTCTGTGACAGCCGAGAGCAAGAACACTCCGAGGCATCAAGGCTCGATCT 1380
Qy 1381 TCACTGGACTCAGAGGATGTTGAAATAGAGTAAACCAAGTTTGTGATGAGCAGCCTTCT 1440
Db 1381 TCACTGGACTCAGAGGATGTTGAAATAGAGTAAACCAAGTTTGTGATGAGCAGCCTTCT 1440
Qy 1441 GCAACACCTGAGAGTGATGCAAGGGTTTCAAGTGGAGCAGAGACAGAAACAAGTTGAC 1500
Db 1441 GCAACACCTGAGAGTGATGCAAGGGTTTCAAGTGGAGCAGAGACAGAAACAAGTTGAC 1500
Qy 1501 CGGTCTCTGTGGCTCAAACTCCGTCCGATGATGATGTTGAGGCGGATGCATCA 1560
Db 1501 CGGTCTCTGTGGCTCAAACTCCGTCCGATGATGATGTTGAGGCGGATGCATCA 1560
Qy 1561 GAAAGCAAGAGATGGCACCCTAGTGGAGTGAAGAAACGAATGAAGACACTAATAAA 1620
Db 1561 GAAAGCAAGAGATGGCACCCTAGTGGAGTGAAGAAACGAATGAAGACACTAATAAA 1620
Qy 1621 CCTCAAACTTCAGATCCCAATGACCGCGAGTAGAATCAGCTCCCAATTAACCGATCCA 1680
Db 1621 CCTCAAACTTCAGATCCCAATGACCGCGAGTAGAATCAGCTCCCAATTAACCGATCCA 1680
Qy 1681 TGAAGTCTGTGCTGACGAGGGTCAAAATGCTTCCAAAGCTCTCTTCTCCAGAGGTA 1740
Db 1681 TGAAGTCTGTGCTGACGAGGGTCAAAATGCTTCCAAAGCTCTCTTCTCCAGAGGTA 1740
Qy 1741 TTGCCCCAAGTTTTTACATATCGAGAAGAACACAGAGAGGAAGAACAAACAAGAA 1800
```

```
Db 1741 TTGCGCAAGTTTTCATATTCGAGAAGAACACAGAGAGGAAGAACAAACAAGAA 1800
Qy 1801 CAAAGATATCAATGGCAGTCTGATCTTAACTTACAGCTAGTTAAACACAGATTGATGAT 1860
Db 1801 CAAAGATATCAATGGCAGTCTGATCTTAACTTACAGCTAGTTAAACACAGATTGATGAT 1860
Qy 1861 CAAAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTAGAAGCTTCAAGCTAATG 1920
Db 1861 CAAAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTAGAAGCTTCAAGCTAATG 1920
Qy 1921 AGTAGAGAAAGAACAGGTTTTTAAACCATACAAAGAGATGTTCCATGGAGCCCAAGAAAGT 1980
Db 1921 AGTAGAGAAAGAACAGGTTTTTAAACCATACAAAGAGATGTTCCATGGAGCCCAAGAAAGT 1980
Qy 1981 AGAATCCTCAACAACATCCTATCATCTGTGGAACAGAAAGATCCCAACCGATGCGG 2040
Db 1981 AGAATCCTCAACAACATCCTATCATCTGTGGAACAGAAAGATCCCAACCGATGCGG 2040
Qy 2041 TTGGAACCTCAAGCTTCCACATGAGACTATTTTCACTGATCTGTTGTTGACTCTG 2100
Db 2041 TTGGAACCTCAAGCTTCCACATGAGACTATTTTCACTGATCTGTTGTTGACTCTG 2100
Qy 2101 TTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTGGAGCCCTTTGATTGTT 2160
Db 2101 TTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTGGAGCCCTTTGATTGTT 2160
Qy 2161 TTCTTGTCTCAGTCTTCTCTGTAACATTTTGACTCTGTTATTTCAACAATCAATAACT 2220
Db 2161 TTCTTGTCTCAGTCTTCTCTGTAACATTTTGACTCTGTTATTTCAACAATCAATAACT 2220
Qy 2221 GTTTAATCTTTTCTTTTCTCA 2240
Db 2221 GTTTAATCTTTTCTTTTCTCA 2240
```

```
RESULT 3
AX506805 LOCUS 1827 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1500 from Patent WO0216655.
ACCESSION AX506805
VERSION AX506805.1 GI:23388042
KEYWORDS
```

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1  
AUTHORS Harper, J.F., Krebs, J., Wang, X. and Zhu, T.  
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use  
JOURNAL Patent: WO 0216655-A 1500 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG (CH)

FEATURES  
source Location/Qualifiers  
1..1827  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"  
BASE COUNT 591 a 397 c 419 g 420 t  
ORIGIN

Query Match 81.1%; Score 1827; DB 6; Length 1827;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 ATGGAGACAAATTCGTCTGGAGAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA 297  
Db 1 ATGGAGACAAATTCGTCTGGAGAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA 60  
Qy 298 ACAAGCAACGTTGAAAGGTGGACTGAGGAAGAACATAATAGATTTCATTGAAGCTTTGAGG 357  
Db 61 ACAAGCAACGTTGAAAGGTGGACTGAGGAAGAACATAATAGATTTCATTGAAGCTTTGAGG 120

QY 358 CTTTATGTTAGACATGGCGAAGATTGAAGAAACATGTAGCAACAAAACTGCTGTCCAG 417  
DB 121 CTTTATGTTAGACATGGCGAAGATTGAAGAAACATGTAGCAACAAAACTGCTGTCCAG 180  
QY 418 ATAAGAAGTCACGCTCAGAAATTTTCTCCAAGTAGAGAAAGGCTGAAGCTAAAGGT 477  
DB 181 ATAAAGAAGTCACGCTCAGAAATTTTCTCCAAGTAGAGAAAGGCTGAAGCTAAAGGT 240  
QY 478 GTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCTTAAGCGTAAACCA 537  
DB 241 GTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCTTAAGCGTAAACCA 300  
QY 538 AACAACTCTTATCTCGAAGACGGGAAGTGGAAACGATCTTATGTCAAAAAAGGCTGTG 597  
DB 301 AACAACTCTTATCTCGAAGACGGGAAGTGGAAACGATCTTATGTCAAAAAAGGCTGTG 360  
QY 598 AATGATGGAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGATGGCCAAATGAA 657  
DB 361 AATGATGGAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGATGGCCAAATGAA 420  
QY 658 GATCGCAACAAATCAAGGCTGAAGGAAACCTCTGAGGAAGACAACTGTTTCAGATTGT 717  
DB 421 GATCGCAACAAATCAAGGCTGAAGGAAACCTCTGAGGAAGACAACTGTTTCAGATTGT 480  
QY 718 TTCACCTCATCAGTATCTCTGCTGTCATCTCTCCATGAATAAAAGTTGTATAGAGATCA 777  
DB 481 TTCACCTCATCAGTATCTCTGCTGTCATCTCTCCATGAATAAAAGTTGTATAGAGATCA 540  
QY 778 AACGCAAGCACTTTCGCGAGTTCTTGCCCTTCACGGGAGAGGAAAGTCAAGATAACAGG 837  
DB 541 AACGCAAGCACTTTCGCGAGTTCTTGCCCTTCACGGGAGAGGAAAGTCAAGATAACAGG 600  
QY 838 GTAAGAAAGAGTCAAACTCAGATTTTCAATGCAAAATCTCTGAAAAACGGTAATAGCAA 897  
DB 601 GTAAGAAAGAGTCAAACTCAGATTTTCAATGCAAAATCTCTGAAAAACGGTAATAGCAA 660  
QY 898 GGCACCTCAGACTTATCGGATGCATATCCCTGCTGCTAGTGCATTTGGGAGCTCAATAACA 957  
DB 661 GGCACCTCAGACTTATCGGATGCATATCCCTGCTGCTAGTGCATTTGGGAGCTCAATAACA 720  
QY 958 AGTTCTCTATCATCATCTCTCCTCAGAGCCAGATAGTCATCCCCACACAGTTGCAGGAGAT 1017  
DB 721 AGTTCTCTATCATCATCTCTCCTCAGAGCCAGATAGTCATCCCCACACAGTTGCAGGAGAT 780  
QY 1018 TATCAGTCGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTAACT 1077  
DB 781 TATCAGTCGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTAACT 840  
QY 1078 GCGCAACTTTCGCTCATCATTTTGGCCCTCCGATTTAGTGGTGGCTCACCTGTTCGA 1137  
DB 841 GCGCAACTTTCGCTCATCATTTTGGCCCTCCGATTTAGTGGTGGCTCACCTGTTCGA 900  
QY 1138 GGAACCTACCTCCGAATCTGGCTGCATGGCCGCGAGCCACTGTTGCAGCTGCTAGTGCT 1197  
DB 901 GGAACCTACCTCCGAATCTGGCTGCATGGCCGCGAGCCACTGTTGCAGCTGCTAGTGCT 960  
QY 1198 TGTGGGCTGCCAATGGATTATTAACCTTTATGTCCTCTTATAGTTTCAAGTGTTCACCT 1257  
DB 961 TGTGGGCTGCCAATGGATTATTAACCTTTATGTCCTCTTATAGTTTCAAGTGTTCACCT 1020  
QY 1258 AGTCATCTCTCATCTACTTTTGGACCATCATGTGATGTAGAGTACAAAAAGCAAGCACT 1317  
DB 1021 AGTCATCTCTCATCTACTTTTGGACCATCATGTGATGTAGAGTACAAAAAGCAAGCACT 1080  
QY 1318 TTACAACATGGTTCTGTGAGAGCCGAGAGCAAGAACACTCCGAGGCATCAAGGCTCGA 1377  
DB 1081 TTACAACATGGTTCTGTGAGAGCCGAGAGCAAGAACACTCCGAGGCATCAAGGCTCGA 1140  
QY 1378 TCTTCACTGGAATCAGAGGATGTTGAAAAATAAGATTAACCAAGTGTGATGAGCAGCT 1437  
DB 1141 TCTTCACTGGAATCAGAGGATGTTGAAAAATAAGATTAACCAAGTGTGATGAGCAGCT 1200

QY 1438 TCTGCAACACCTGAGAGTGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAACAAGTT 1497  
DB 1201 TCTGCAACACCTGAGAGTGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAACAAGTT 1260  
QY 1498 GACCGGTCTCTGCTCGCTCAAAACACTCCGTCAGTAGTGTGATGTTGGGCGGATGCA 1557  
DB 1261 GACCGGTCTCTGCTCGCTCAAAACACTCCGTCAGTAGTGTGATGTTGGGCGGATGCA 1320  
QY 1558 TCAGAAAGGCAAGAGGATGGCACCAATGGTGGAGTGAAGAAACCAATGAAGACACTAAT 1617  
DB 1321 TCAGAAAGGCAAGAGGATGGCACCAATGGTGGAGTGAAGAAACCAATGAAGACACTAAT 1380  
QY 1618 AAACCTCAAACTTCAGAGTCCAATGCACGCCGAGTAGAATCAGCTCAATATAACCGAT 1677  
DB 1381 AAACCTCAAACTTCAGAGTCCAATGCACGCCGAGTAGAATCAGCTCAATATAACCGAT 1440  
QY 1678 CCATGGAAGTCTCTGCTGACGAGGTCGAAATTCGCTTCCAAGCTCTCTTCTCCAGAG 1737  
DB 1441 CCATGGAAGTCTCTGCTGACGAGGTCGAAATTCGCTTCCAAGCTCTCTTCTCCAGAG 1500  
QY 1738 GTATTCCGCAAAAGTTTTACATATCGAAGAAACACAGAGAGGAAGAAACAACAACAA 1797  
DB 1501 GTATTCCGCAAAAGTTTTACATATCGAAGAAACACAGAGAGGAAGAAACAACAACAA 1560  
QY 1798 GAACAAAGATATCCAATGGCATTGATCTTTAACTTCACAGCTCAGTTAAACACCACTTGT 1857  
DB 1561 GAACAAAGATATCCAATGGCATTGATCTTTAACTTCACAGCTCAGTTAAACACCACTTGT 1620  
QY 1858 GATCAAGAGGAGAGAGAAACACAGGATTTCTTGGATCGGATAGATGCTTCAAGCTA 1917  
DB 1621 GATCAAGAGGAGAGAGAAACACAGGATTTCTTGGATCGGATAGATGCTTCAAGCTA 1680  
QY 1918 ATCAGTAGAGAGAGAAACAGGTTTTTAAACCATACAAAGAGTGTCCATGGAGCCAAAGAA 1977  
DB 1681 ATCAGTAGAGAGAGAAACAGGTTTTTAAACCATACAAAGAGTGTCCATGGAGCCAAAGAA 1740  
QY 1978 AGTAGAATCTCAACAAACATCTCATCTATCTATGTGGAAACAGAAAGATCCCAACCGATG 2037  
DB 1741 AGTAGAATCTCAACAAACATCTCATCTATCTATGTGGAAACAGAAAGATCCCAACCGATG 1800  
QY 2038 CGGTTGGAACCTCAAGCTTCCACATCA 2064  
DB 1801 CGGTTGGAACCTCAAGCTTCCACATCA 1827

## RESULT 4

AY050961  
LOCUS Arabidopsis thaliana At2g46830 mRNA linear PLN 23-SEP-2002  
DEFINITION Arabidopsis thaliana At2g46830 mRNA sequence.  
ACCESSION AY050961  
VERSION AY050961.1 GI:15293054  
KEYWORDS FLI CDNA.

## SOURCE

Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

1 (bases 1 to 2684)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,  
Quach, H.L., Tang, C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C.,  
Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,  
Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,  
Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,  
Ecker, J.R. and Theologis, A.

## AUTHORS

Arabidopsis Full length cDNA Clones  
Unpublished  
2 (bases 1 to 2684)  
Yamada, K., Liu, S.X., Pham, P.K., Banh, J., Banno, F., Dale, J.M.,  
Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L.,  
Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C.,

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

# Direct Submission

Submitted (01-AUG-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

## REFERENCE

### AUTHORS

Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Yuen, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

# Direct Submission

Submitted (23-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

## COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

## FEATURES

### source

1. 2684  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAFL05-20-N17 (R10389)"  
/note="This clone is in a modified pluescript vector (Lambda Zap) as a XhoI/SstI insert.  
ecotype: Columbia"

### gene

1. 2684  
/gene="At2g46830"  
66. 2478  
/gene="At2g46830"

### misc\_feature

/note="This may be a potential product of a partially processed transcript corresponding to gene At2g46830."  
/evidence=experimental

### misc\_difference

1772  
/gene="At2g46830"  
/note="compared to genomic sequence"  
/replace="C"

### misc\_difference

2655  
/gene="At2g46830"  
/note="not present in genomic sequence"  
2668  
/gene="At2g46830"

### misc\_feature

/note="artifact within poly A tail"  
BASE COUNT 831 a 525 c 561 g 767 t  
ORIGIN

Query Match 80.1%; Score 1805; DB 8; Length 2684;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1808; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

## QY

441 TTCTCCAGGTAGAGAAAGAGGCTGAAGCTTAAGTGTAGCTATGGTCAAGCGCTAGA 500

## Db

855 TCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTTAAGTGTAGCTATGGTCAAGCGCTAGA 914

QY 501 CATAGCTATTCTCTCCACGGCCTTAAGCGTAAACCAAACTCTTATCTCTCGAAAGAC 560  
Db 915 CATAGCTATTCTCTCCACGGCCTTAAGCGTAAACCAAACTCTTATCTCTCGAAAGAC 974  
QY 561 GGGAAAGTGGAAACGATCTTATGTCAAAAACGGGTGTGAATGATGAAAAGAGTCTCTTGG 620  
Db 975 GGGAAAGTGGAAACGATCTTATGTCAAAAACGGGTGTGAATGATGAAAAGAGTCTCTTGG 1034  
QY 621 ATCAGAAAAAGTGTGCGATCTGAGATGGCCAAATGAAGATCGAACAACTCAAGACCTGA 680  
Db 1035 ATCAGAAAAAGTGTGCGATCTGAGATGGCCAAATGAAGATCGAACAACTCAAGACCTGA 1094  
QY 681 AGAAAAAATCTGCGAGGAAGACAACTGTTCAGATGTGTTTCACTCATCATGATCTCTCTGC 740  
Db 1095 AGAAAAAATCTGCGAGGAAGACAACTGTTCAGATGTGTTTCACTCATCATGATCTCTCTGC 1154  
QY 741 TGCATCTCCATGAATAAAAGTTGTATAGAGACATCAAAACCAAGCACTCTTCCGCGAGTT 800  
Db 1155 TGCATCTCCATGAATAAAAGTTGTATAGAGACATCAAAACCAAGCACTCTTCCGCGAGTT 1214  
QY 801 CTTCCTTTCAGGGAAGAGGAACTCAGATAACAGGGTAAGAAAGAGTCAAACTCAGA 860  
Db 1215 CTTCCTTTCAGGGAAGAGGAACTCAGATAACAGGGTAAGAAAGAGTCAAACTCAGA 1274  
QY 861 TTTGAATGCAAAATCTCTGAAAAACGGTAATAGCAAGGACCTCAGACTTATCCGATGCA 920  
Db 1275 TTTGAATGCAAAATCTCTGAAAAACGGTAATAGCAAGGACCTCAGACTTATCCGATGCA 1334  
QY 921 TATCCCTGTGCTAGTGCATTTGGGAGCTCAATAAACAAAGTCTCTATCACATCTCTCTTC 980  
Db 1335 TATCCCTGTGCTAGTGCATTTGGGAGCTCAATAAACAAAGTCTCTATCACATCTCTCTTC 1394  
QY 981 AGAGCCAGATAGTATCATCCCAACAGTTGCAGGAGATTATCAGTCGTTTCTTAATCATAT 1040  
Db 1395 AGAGCCAGATAGTATCATCCCAACAGTTGCAGGAGATTATCAGTCGTTTCTTAATCATAT 1454  
QY 1041 AATGTCAACCTTTTACAAAACACGGCTCTTTATACCTCCGCAACTTTCGCTCATCATTT 1100  
Db 1455 AATGTCAACCTTTTACAAAACACGGCTCTTTATACCTCCGCAACTTTCGCTCATCATTT 1514  
QY 1101 TTGGCTCCCGATTCTAGTGGTGGTCCACCTGTTCCAGGGAACCTCAGCTCCGAATCTGGC 1160  
Db 1515 TTGGCTCCCGATTCTAGTGGTGGTCCACCTGTTCCAGGGAACCTCAGCTCCGAATCTGGC 1574  
QY 1161 TGCATGGCCGACGACCTGTTCAGCTGTAGTGTGTTGGGCTGCCAATGATTAAT 1220  
Db 1575 TGCATGGCCGACGACCTGTTCAGCTGTAGTGTGTTGGGCTGCCAATGATTAAT 1634  
QY 1221 ACCTTTATGTGCTCTCTTATAGTTTCACTAGTATCATCTCCATCTACTCTTTTGG 1280  
Db 1635 ACCTTTATGTGCTCTCTTATAGTTTCACTAGTATCATCTCCATCTACTCTTTTGG 1694  
QY 1281 ACCATCATGTAGTATAGTATACAAAAGCAAGCACTTTTCAACATGGTCTGTGCGAGAG 1340  
Db 1695 ACCATCATGTAGTATAGTATACAAAAGCAAGCACTTTTCAACATGGTCTGTGCGAGAG 1754  
QY 1341 CCGAGAGCAAGAACACTCCGAGGATCAAGGCTCGATCTTCACTGACATCAGAGGATGT 1400  
Db 1755 CCGAGAGCAAGAACACTTCGAGGATCAAGGCTCGATCTTCACTGACATCAGAGGATGT 1814  
QY 1401 TGAATAAAGAGTAAACCAAGTTGTCTAGTACGAGCTCTTGCACACCTCGAGAGTATGC 1460  
Db 1815 TGAATAAAGAGTAAACCAAGTTGTCTAGTACGAGCTCTTGCACACCTCGAGAGTATGC 1874  
QY 1461 AAAGGCTTCAGATGGAGCAGGACACAAAACAAAGTTGACCGTCTCTGTGTGGCTCAAA 1520  
Db 1875 AAAGGCTTCAGATGGAGCAGGACACAAAACAAAGTTGACCGTCTCTGTGTGGCTCAAA 1934  
QY 1521 CACTCCGTCTGAGTGTAGTATGTTGAGGCGGATGATCAGAAAGCAAGAGATGCGCAC 1580  
Db 1935 CACTCCGTCTGAGTGTAGTATGTTGAGGCGGATGATCAGAAAGCAAGAGATGCGCAC 1994  
QY 1581 CAAATGTGTAGGTGAAAGAAAACGAATGAAGACACTAATAAAACCTTCAAACTTCAGATCCAA 1640



Db 421 CAGACTTATCCGATGCTATCCCTGTGCTAGTGCATTGGGAGCTCAATAACAAGTTCT 480  
Qy 964 CTATCAGATCCTCCTTCAGACCCAGATAGTCAATCCACACAGTTCAGGAGATTATCAG 1023  
Db 481 CTATCAGATCCTCCTTCAGACCCAGATAGTCAATCCACACAGTTCAGGAGATTATCAG 540  
Qy 1024 TCGTTTCTTAAATCATATAATGTCACCCCTTTTACAAACACCCGGCTCTTTTATCTGCCGCA 1083  
Db 541 TCGTTTCTTAAATCATATAATGTCACCCCTTTTACAAACACCCGGCTCTTTTATCTGCCGCA 600  
Qy 1084 ACTTTCGGCTCATCATTTTGGCCCTCCCGATTTCTAGTGTGCTCACTGTTCCAGGAAC 1143  
Db 601 ACTTTCGGCTCATCATTTTGGCCCTCCCGATTTCTAGTGTGCTCACTGTTCCAGGAAC 660  
Qy 1144 TCACCTCCGAATCTGGCTGCCATGCGCCGACGCACTGTTGAGCTGTAGTGTGCTGG 1203  
Db 661 TCACCTCCGAATCTGGCTGCCATGCGCCGACGCACTGTTGAGCTGTAGTGTGCTGG 720  
Qy 1204 GCTGCCAATGGATTATTACCTTTTATGTGCTCTCTTAGTTTCTAGTGTGCTTCACTAGTCAT 1263  
Db 721 GCTGCCAATGGATTATTACCTTTTATGTGCTCTCTTAGTTTCTAGTGTGCTTCACTAGTCAT 780  
Qy 1264 CTTCCATCTACTTTTGGACCATCATGTGATGTAGTACACAAAGCAAGCACTTTACAA 1323  
Db 781 CTTCCATCTACTTTTGGACCATCATGTGATGTAGTACACAAAGCAAGCACTTTACAA 840  
Qy 1324 CATGGTTCTGTGACAGCCGAGACAAAGCACTCCGAGGATCAAAAGGCTCGATCTTCA 1383  
Db 841 CATGGTTCTGTGACAGCCGAGACAAAGCACTCCGAGGATCAAAAGGCTCGATCTTCA 900  
Qy 1384 CTGGAATCAGAGGATGTTGAAATTAAGATTAACACGTTTGTATGACGAGCTTTCTGCA 1443  
Db 901 CTGGAATCAGAGGATGTTGAAATTAAGATTAACACGTTTGTATGACGAGCTTTCTGCA 960  
Qy 1444 ACACCTCAGAGTGTGCAAGGTTTCAGATGAGGACGAGACAGAAACAAAGTTGACCGG 1503  
Db 961 ACACCTCAGAGTGTGCAAGGTTTCAGATGAGGACGAGACAGAAACAAAGTTGACCGG 1020  
Qy 1504 TCCTCGTGTGGCTCAAACTCCGTCGAGTAGTATGATGTTGAGGCGGATGCATCAGAA 1563  
Db 1021 TCCTCGTGTGGCTCAAACTCCGTCGAGTAGTATGATGTTGAGGCGGATGCATCAGAA 1080  
Qy 1564 AGGCAAGAGATGCAATGTTGAGTGAAGAAACGAATGAAGACACTAATAAACCT 1623  
Db 1081 AGGCAAGAGATGCAATGTTGAGTGAAGAAACGAATGAAGACACTAATAAACCT 1140  
Qy 1624 CAAACTTCAGATCCCAATGCAAGCCGAGTAGATCAGCTTCAATATACCGATCCATGG 1683  
Db 1141 CAAACTTCAGATCCCAATGCAAGCCGAGTAGATCAGCTTCAATATACCGATCCATGG 1200  
Qy 1684 AAGTCTGTGTCTGACGAGGTCGAATTCGCTTCCAACTCTCTTCTCCAGAGAGGTATTG 1743  
Db 1201 AAGTCTGTGTCTGACGAGGTCGAATTCGCTTCCAACTCTCTTCTCCAGAGAGGTATTG 1260  
Qy 1744 CCGCAAGTTTTCATATCAGAGAAGAACACAGAGAGGAAGAACAAACAAACAAAGAACAA 1803  
Db 1261 CCGCAAGTTTTCATATCAGAGAAGAACACAGAGAGGAAGAACAAACAAACAAAGAACAA 1320  
Qy 1804 AGATATCCCAATGGCACTTGTATCTTAACTTCACAGCTCAGTTTAAACAGTTGATGATCAA 1863  
Db 1321 AGATATCCCAATGGCACTTGTATCTTAACTTCACAGCTCAGTTTAAACAGTTGATGATCAA 1380  
Qy 1864 GAGGAGAAGAAACACAGGATTTCTTGGAAATCGGATTAGATGCTTCAAAAGCTTAATGAGT 1923  
Db 1381 GAGGAGAAGAAACACAGGATTTCTTGGAAATCGGATTAGATGCTTCAAAAGCTTAATGAGT 1440  
Qy 1924 AGAGGAAGAACAGGTTTAAACCATACAAAGATGTTCCATGGAAGCCAAAGAAAGTAGA 1983  
Db 1441 AGAGGAAGAACAGGTTTAAACCATACAAAGATGTTCCATGGAAGCCAAAGAAAGTAGA 1500  
Qy 1984 ATCCTCAACAAACATCTATCTATCTTGTGGAACAGAAAGATCCCAACCGATGCGGTTG 2043

1501 ATCCTCAACAACATCTATCATTCATGTGGACAGAAAGATCCAAACGGATCGGTTG 1560  
Qy 2044 GAAACTCAAGCTTCCACATGAGACTATTTTTCATCTGATCTGTTTGTGA 2095  
Db 1561 GAAACTCAAGCTTCCACATGAGACTATTTTTCATCTGATCTGTTTGTGA 1612

RESULT 6  
BT001105 1612 bp mRNA linear PLN 07-NOV-2002  
LOCUS Arabidopsis thaliana clone C105127 (D) putative MYB-related  
DEFINITION transcription factor CCA1 (At2g46830) mRNA, complete cds.  
ACCESSION BT001105  
VERSION BT001105.1 GI:24762204  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 1612)  
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,  
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,  
Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M.,  
Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W.,  
Ecker,J.R. and Theologis,A.  
TITLE Arabidopsis Open Reading Frame (ORF) Clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1612)  
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,  
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,  
Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M.,  
Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W.,  
Ecker,J.R. and Theologis,A.  
TITLE Direct Submission  
JOURNAL Submitted (07-NOV-2002) Plant Gene Expression Center, 800 Buchanan  
COMMENT Annotation based on July 2002 version of the Arabidopsis genome  
submitted to Genbank.

FEATURES  
source  
1..1612  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="C105127 (D)"  
/notes="This clone is in pUNI 51."  
ecotype: Columbia  
1..1612  
/gene="At2g46830"  
1..1581  
/gene="At2g46830"  
/codon\_start=1  
/evidence=experimental  
/product="putative MYB-related transcription factor CCA1"  
/protein\_id="AAN64169.1"  
/db\_xref="GI:24762205"  
/translation="MCQALDIALPPPPRRKPNPNYPRKTSGLTMSKTKVNDGKES  
LGSEKVSHPHEMANEDRQKPEKTLQEDNCDFTHQYLSAASAMNKSCIEITSNAST  
REFLEPSREGQNNVRKESNDLNKSLNENQGPQTPYPMHPIVLPVLSGITS  
LSHPSPEDSHPTVAGDYQSPFNHIMSTLQTPALYTAATFASFPDSSGSPVP  
GNSPNLAAMAATVAASAAMWANGLLPLCAPLSSGGFTSHPSFTGSPCDVEYTKA  
STLQHSVQSRQEHSEASKARSLDSEDEVENKSPVCHQPSATPESDAKGDGADG  
RKQVDRSSCSNTSPSSDDVEADASERQEDNGTEVKETNEDNKPTQSSNARRSI  
SNSITDPKSVSDPSRIFAQALFQVLFQSFYTREREHEEQOQOEPYPMALDLNF  
TQALTVPDQOEEKRNTGFLIGLDASKLMSRGTGPKPKRCSMEAKESILNNPIL  
HVEOKDPKRMRLTQAST"  
1582..1612  
/gene="At2g46830"  
505 a 366 c 360 g 381 t

3'UTR  
BASE COUNT 505 a 366 c 360 g 381 t  
ORIGIN

Query Match 71.5%; Score 1612; DB 8; Length 1612;



Best Local Similarity 100.0%; Pred. No. 0; Matches 1612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	484	ATGGTCAAGCGCTAGACATAGCTATTCCTCTCCACGGCCTTAAGCGTAAACCAACCAAT	543
Db	1	ATGGGTCAAGCGCTAGACATAGCTATTCCTCTCCACGGCCTTAAGCGTAAACCAACCAAT	60
QY	544	CCTTATCTCGAAGACGGGAAGTGAACGATCCTTTATGTCAAAAACGGGTGTGAATGAT	603
Db	61	CCTTATCTCGAAGACGGGAAGTGAACGATCCTTTATGTCAAAAACGGGTGTGAATGAT	120
QY	604	GGAAAGAGTCCCTTGGATCAGAAAAGTGTGCGATCCTGAGATGGCCCAATGAAGTCGA	663
Db	121	GGAAAGAGTCCCTTGGATCAGAAAAGTGTGCGATCCTGAGATGGCCCAATGAAGTCGA	180
QY	664	CAACAATCAAGCCTGAAGAGAAAACTCTGCAGGAAGACAACTGTTCAGATTGTTTCACT	723
Db	181	CAACAATCAAGCCTGAAGAGAAAACTCTGCAGGAAGACAACTGTTCAGATTGTTTCACT	240
QY	724	CATCAGTATCTCTGTGTCATCCTCCATGAATAAAAGTTGTATAGAGACATCAAAACGCA	783
Db	241	CATCAGTATCTCTGTGTCATCCTCCATGAATAAAAGTTGTATAGAGACATCAAAACGCA	300
QY	784	AGCATTTCGCGAGTCTTGCTTTCACGGGAAGGAAGTCAGAAATAACAGGGTAAGA	843
Db	301	AGCATTTCGCGAGTCTTGCTTTCACGGGAAGGAAGTCAGAAATAACAGGGTAAGA	360
QY	844	AAGGAGTCAAACTCAGATTGTAATCAAAATCTCTGGAAAACGGTAATGAGCAAGCACT	903
Db	361	AAGGAGTCAAACTCAGATTGTAATCAAAATCTCTGGAAAACGGTAATGAGCAAGCACT	420
QY	904	CAGACTTATCGATGCATATCCCTGTGTAGTGCCATTGGGGAGTCAATTAACAAGTTCT	963
Db	421	CAGACTTATCGATGCATATCCCTGTGTAGTGCCATTGGGGAGTCAATTAACAAGTTCT	480
QY	964	CTATCACATCTCTTCAGAGCCAGATAGTATCCCCACAGTTGCGAGGATTTATCAG	1023
Db	481	CTATCACATCTCTTCAGAGCCAGATAGTATCCCCACAGTTGCGAGGATTTATCAG	540
QY	1024	TCGTTTCTCAATCATATATATGTCAACCCCTTTTACAAAACCGGCTCTTTATCTGCCGCA	1083
Db	541	TCGTTTCTCAATCATATATATGTCAACCCCTTTTACAAAACCGGCTCTTTATCTGCCGCA	600
QY	1084	ACTTTCGCTCATCATTTTGGCCTCCGATTTCTAGTGTGGTCACTGTTCAGGGGAAC	1143
Db	601	ACTTTCGCTCATCATTTTGGCCTCCGATTTCTAGTGTGGTCACTGTTCAGGGGAAC	660
QY	1144	TCACCTCCGAATCTGGCTGCCATGGCCGACGCTGTTGAGCTGTAGTGTGGTGG	1203
Db	661	TCACCTCCGAATCTGGCTGCCATGGCCGACGCTGTTGAGCTGTAGTGTGGTGG	720
QY	1204	GCTGCCAATGGATTATTAACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGTCAT	1263
Db	721	GCTGCCAATGGATTATTAACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGTCAT	780
QY	1264	CCTCATCTACTTTTGGACCATCATGTGATGTAGAGTACAAAAGCAAGCACTTTACAA	1323
Db	781	CCTCATCTACTTTTGGACCATCATGTGATGTAGAGTACAAAAGCAAGCACTTTACAA	840
QY	1324	CATGGTCTGTGCAGAGCCGAGAGCAAGAACACTCCGAGGATCAAAAGGCTCGATCTTCA	1383
Db	841	CATGGTCTGTGCAGAGCCGAGAGCAAGAACACTCCGAGGATCAAAAGGCTCGATCTTCA	900
QY	1384	CTGACTCAGAGGATGTTGAAAATAAGAGTAAACCCAGTTTGTCTATGACAGCCTTCTGCA	1443
Db	901	CTGACTCAGAGGATGTTGAAAATAAGAGTAAACCCAGTTTGTCTATGACAGCCTTCTGCA	960
QY	1444	ACACTCAGAGTGTGCAAGGGTTTCAGATGGAGCAGGAGACAGAAAACAAGTTGACCGG	1503
Db	961	ACACTCAGAGTGTGCAAGGGTTTCAGATGGAGCAGGAGACAGAAAACAAGTTGACCGG	1020
QY	1504	TCCTCGTGTGGCTCAAAACACTCCCTCGAGTGTAGTATGTTGAGGGGATGCATCAGAA	1563

Db	1021	TCCTCGTGTGGCTCAAAACACTCCGTCGAGTAGTAGTATGATGTTTGAGCGGATGCATCAGAA	1080
QY	1564	AGCCAAGAGTAGGGCACCAATGCTGAGGTGAAAGAAACCAAGTGAAGACACTTAATAAACCT	1623
Db	1081	AGCCAAGAGTAGGGCACCAATGCTGAGGTGAAAGAAACCAAGTGAAGACACTTAATAAACCT	1140
QY	1624	CAAACTTCAGAGTCCAAATGCACGCCGACGTAGAATCAGCTCCTCAATATAACCGATCCATGG	1683
Db	1141	CAAACTTCAGAGTCCAAATGCACGCCGACGTAGAATCAGCTCCTCAATATAACCGATCCATGG	1200
QY	1684	AGTCTGTGTCTGACGAGGGTCAATTGCCTTTCCAAAGCTCTCTTCCAGAGAGGTATTG	1743
Db	1201	AGTCTGTGTCTGACGAGGGTCAATTGCCTTTCCAAAGCTCTCTTCCAGAGAGGTATTG	1260
QY	1744	CGCAAAAGTTTTCACATATCGAGAAGAACACAGAGAGGAAGAACAAACAACAAGAACAA	1803
Db	1261	CGCAAAAGTTTTCACATATCGAGAAGAACACAGAGAGGAAGAACAAACAACAAGAACAA	1320
QY	1804	AGATATCCAATGGCACTTTGATCTTAACCTTCACAGCTCAGTTAACACAGTTGATGATCAA	1863
Db	1321	AGATATCCAATGGCACTTTGATCTTAACCTTCACAGCTCAGTTAACACAGTTGATGATCAA	1380
QY	1864	GAGGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAGCTTAATGAGT	1923
Db	1381	GAGGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAGCTTAATGAGT	1440
QY	1924	AGAGGAAGAACAGGTTTAAACCATACAAAAGATGTTCCATCGAAGCCAAAGAAAGTAGA	1983
Db	1441	AGAGGAAGAACAGGTTTAAACCATACAAAAGATGTTCCATCGAAGCCAAAGAAAGTAGA	1500
QY	1984	ATCCTCAACAAACATCTTATCATTTGGAACAGAAAGATCCCAACCGATCGCGTTG	2043
Db	1501	ATCCTCAACAAACATCTTATCATTTGGAACAGAAAGATCCCAACCGATCGCGTTG	1560
QY	2044	GAACTCAAGCTTCACATGAGACTCTATTTTCATCTGATCTGTTGTTGTA	2095
Db	1561	GAACTCAAGCTTCACATGAGACTCTATTTTCATCTGATCTGTTGTTGTA	1612
RESULT 7			
BD107660			
LOCUS			
DEFINITION			
Phytochrome regulated transcription factor for control of higher			
plant development.			
ACCESSION			
BD107660			
VERSION			
BD107660.1 GI:23202478			
KEYWORDS			
JP 2002501381-A/1.			
SOURCE			
unidentified			
ORGANISM			
unclassified.			
REFERENCE			
1 (bases 1 to 4344)			
Tobin,R., Sun,L. and Wang,Z.Y.			
Phytochrome regulated transcription factor for control of higher			
plant development			
Patent: JP 2002501381-A 1 15-JAN-2002;			
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA			
COMMENT			
OS Unidentified			
PN JP 2002501381-A/1			
PD 15-JAN-2002			
PF 17-APR-1998 JP 1998546171			
PR 18-APR-1997 US 08/843572			
PI ROLAND TOBIN,LIN SUN,ZHI YONG WANG			
PC C12N15/00,C12N15/09,C12N15/29,C12N15/82,A01H1/00,A01H3/00, PC			
A01H5/00			
CC Strandedness: Single;			
CC Topology: Linear;			
CC Phytochrome regulated transcription factor for control of CC			
higher plant			
CC development			
FH Key			
Location/Qualifiers			
I. .4344			
FT source			
/organism='Unidentified'.			
FT Location/Qualifiers			

[illegible]

[illegible]





repeat\_region /rpt\_family="AT\_rich"  
28087. .28145  
/rpt\_family=" (TAAAA)n"

Query Match		67.7%;	Score 1526.6;	DB 8;	Length 68415;
Best Local Similarity		87.8%;	Pred. No. 0;		
Matches 1803;		Conservative 0;	Mismatches 9;	Indels 241;	Gaps 3;
Qy	441	TTTCTCCAAAGGTAGAGAAAGAGGCTGAAGCTTAAAGGTGTAGCTATGGGTCAAGCGCTAGA	500		
Db	46061	TCTCTCAGGTAGAGAAAGAGGCTGAAGCTTAAAGGTGTAGCTATGGGTCAAGCGCTAGA	46120		
Qy	501	CATAGCTATTCTCTCCACCGCCTTAAGCGTAAACCAACCAATCCTTATCTCTCGAAAGAC	560		
Db	46121	CATAGCTATTCTCTCCACCGCCTTAAGCGTAAACCAACCAATCCTTATCTCTCGAAAGAC	46180		
Qy	561	GGGAAGTGAACGATCCTTATGTCAAAAACGGGTGTGAATCATGGAAAAGAGTCCCTTGG	620		
Db	46181	GGGAAGTGAACGATCCTTATGTCAAAAACGGGTGTGAATCATGGAAAAGAGTCCCTTGG	46240		
Qy	621	ATCAGAAAAAGTGTGCAATCCTG-----	643		
Db	46241	ATCAGAAAAAGTGTGCAATCCTGAGGTGATTTTCATGGTCATATGGCATCTTTTTCAGT	46300		
Qy	644	-----	653		
Db	46301	GTGTCACATTTGCTCTCATGTTATTAATACAGATTGTGTGCTTCGTTTATAGATGGCCAA	46360		
Qy	654	TGAAGATCGCAACAATCAAAAGCCTGAAGAGAAAACCTCTGCAGGAGACAACCTGTTTACA	713		
Db	46361	TGAAGATCGCAACAATCAAAAGCCTGAAGAGAAAACCTCTGCAGGAGACAACCTGTTTACA	46420		
Qy	714	TTGTTTCACTCATCAGTATCTCTGTGTCATCTCTCATGAATATAAAGTTGTTATAGAGAC	773		
Db	46421	TTGTTTCACTCATCAGTATCTCTGTGTCATCTCTCATGAATATAAAGTTGTTATAGAGAC	46480		
Qy	774	ATCAAAACCGACGATTTCCCGGAGTCTTCCCTTTCACGGGAAG-----	817		
Db	46481	ATCAAAACCGACGATTTCCCGGAGTCTTCCCTTTCACGGGAAGAGGTAAAAACAATCT	46540		
Qy	818	-----	817		
Db	46541	TTTCATTGCTATTTCAGGTTTTAAGCAGATTAGTACTTTTCATGAACCTAAACCGTGGGG	46600		
Qy	818	-----AGGAAGTCAGATAACAGGGTAAGAAAGGAGTCAAACTCAGATTTCGAATGCA	870		
Db	46601	GAATAACAGGAAGTCAGATAACAGGGTAAGAAAGGAGTCAAACTCAGATTTCGAATGCA	46660		
Qy	871	AAATCTCTGNAACCGTAAATGACCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG	930		
Db	46661	AAATCTCTGNAACCGTAAATGACCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG	46720		
Qy	931	CTAGTGCATTGGGGAGCTCAATAACAAAGTTCTCTATCACATCTCTTCAGAGCCAGAT	990		
Db	46721	CTAGTGCATTGGGGAGCTCAATAACAAAGTTCTCTATCACATCTCTTCAGAGCCAGAT	46780		
Qy	991	AGTCATCCCCACA CAGTTGCAAGAGATTATCAGTCGTTTCCATAATATAATGTCAACC	1050		
Db	46781	AGTCATCCCCACA CAGTTGCAAGAGATTATCAGTCGTTTCCATAATATAATGTCAACC	46840		
Qy	1051	CTTTTACAAACACCGGCTTTTATCTGCGCAACTTTTCGCTCATCATTTTGGCCCTCC	1110		
Db	46841	CTTTTACAAACACCGGCTTTTATCTGCGCAACTTTTCGCTCATCATTTTGGCCCTCC	46900		
Qy	1111	GATTCTAGTGGTGCTCACCTGTTTCCAGGGAACCTCACCTCCGAATCTGGCTGCCATGSCC	1170		
Db	46901	GATTCTAGTGGTGCTCACCTGTTTCCAGGGAACCTCACCTCCGAATCTGGCTGCCATGSCC	46960		
Qy	1171	GCAGCCACTGTTGCAGCTGCTAGTGTGTTGGGCTGCCAATGATTATTACCTTTATGT	1230		
Db	46961	GCAGCCACTGTTGCAGCTGCTAGTGTGTTGGGCTGCCAATGATTATTACCTTTATGT	47020		
Qy	1231	GCTCCTCTTAGTTTCAGGTGGTTTTCAGTACTATCCTCCTACTTCTTTTGACCAACATCATGT	1290		

Db	47021	 GCTCCTCTTAGTTTCAGGTGGTTTTCAGTACTCTCTCATCTACTCTTTTGGACCATCATGT	47080
Qy	1291	GATGTAGAGTACACAAAAGCAAGCACTTTACAACTAGTGTCTGTGTCAGAGCCGAGAGCAA	1350
Db	47081	GATGTAGAGTACACAAAAGCAAGCACTTTACAACTAGTGTCTGTGTCAGAGCCGAGAGCAA	47140
Qy	1351	GAACACTCCGAGGCATCAAAGGCTCGATCTTCTCACTGGACTCAGAGGATGTTGAAAAATAAG	1410
Db	47141	GAACACTCCGAGGCATCAAAGGCTCGATCTTCTCACTGGACTCAGAGGATGTTGAAAAATAAG	47200
Qy	1411	AGTAAACCAAGTTTGTCTATGAGCAGCCTTCTGCAACACCTGAGAGTGTATGCAAGGGTTCA	1470
Db	47201	AGTAAACCAAGTTTGTCTATGAGCAGCCTTCTGCAACACCTGAGAGTGTATGCAAGGGTTCA	47260
Qy	1471	GATGAGCAGGAGACAGAAAAACAAGTTGACCGTCTCGTGTGGCTCAAACTCCGTCG	1530
Db	47261	GATGAGCAGGAGACAGAAAAACAAGTTGACCGTCTCGTGTGGCTCAAACTCCGTCG	47320
Qy	1531	AGTAGTGATGATGTTTGAGCGGATGCATCAGAAAGGCAAGAGATGGCACCAATGGTGAG	1590
Db	47321	AGTAGTGATGATGTTTGAGCGGATGCATCAGAAAGGCAAGAGATGGCACCAATGGTGAG	47380
Qy	1591	GTGAAAGAAA CGAATGAAGACACTAATAAACTTCAAACTTTCAGAGTCCAAATGCCACGCC	1650
Db	47381	GTGAAAGAAA CGAATGAAGACACTAATAAACTTCAAACTTTCAGAGTCCAAATGCCACGCC	47440
Qy	1651	AGTAGAATCAGCTCCAAATATAACCGATCCATCGAAGTCTGTGCTGACG----- 1699	
Db	47441	AGTAGAATCAGCTCCAAATATAACCGATCCATCGAAGTCTGTGCTGACGAGGTACTTACT	47500
Qy	1700	----- 1699	
Db	47501	TGGACTAAAGATCAACTCTCTTTTATTTCAAATCATTTTCTCATATAAATATTTGACATTC	47560
Qy	1700	AGGTCGAAATTCGCTTCCAAAGCTCTTCTTCAGAGAGGTATTCGCGCAAAAGTTTACAT	1759
Db	47561	AGGTCGAAATTCGCTTCCAAAGCTCTTCTTCAGAGAGGTATTCGCGCAAAAGTTTACAT	47620
Qy	1760	ATCGAGAAGAACACAGAGAGGAAGAACACAAACAAAGAACAAAGATATCCAAATGGCAC	1819
Db	47621	ATCGAGAAGAACACAGAGAGGAAGAACACAAACAAAGAACAAAGATATCCAAATGGCAC	47680
Qy	1820	TTGATCTTTAACTTCCACAGCTCAGTTTAAACACCGTTGATGATCAAGAGGAGAGAAACA	1879
Db	47681	TTGATCTTTAACTTCCACAGCTCAGTTTAAACACCGTTGATGATCAAGAGGAGAGAAACA	47740
Qy	1880	CAGGATTTCTTGGAAATCGGATTAGATGCTTCAAAGCTAATGAGTAGAGGAGAAACAGGTT	1939
Db	47741	CAGGATTTCTTGGAAATCGGATTAGATGCTTCAAAGCTAATGAGTAGAGGAGAAACAGGTT	47800
Qy	1940	TTAAACCATACAAAAGATGTTCCATGGAGCCAAAGAAAGTAAATCTCTCAACAAATC	1999
Db	47801	TTAAACCATACAAAAGATGTTCCATGGAGCCAAAGAAAGTAAATCTCTCAACAAATC	47860
Qy	2000	CTATCATTTCTGTGGAACAGAAAGATCCCAAAACGGATCGGTTGGAAACTCAAGCTTCCA	2059
Db	47861	CTATCATTTCTGTGGAACAGAAAGATCCCAAAACGGATCGGTTGGAAACTCAAGCTTCCA	47920
Qy	2060	CATGAGACTCTATTTTTCATCTGATCTGTGTTTGTACTCTGTTTTTAAAGTTTTCAAGACC	2119
Db	47921	CATGAGACTCTATTTTTCATCTGATCTGTGTTTGTACTCTGTTTTTAAAGTTTTCAAGACC	47980
Qy	2120	ACTGCTACATTTCTTTTCTTTTCAGGCTTTGTTGTTTCTCTGTTTCTTGTCTTCTTCTTC	2179
Db	47981	ACTGCTACATTTCTTTTCTTTTCAGGCTTTGTTGTTTCTCTGTTTCTTGTCTTCTTCTTC	48040
Qy	2180	CTGTAAACATTTGACTCTGTATTTTCAACAAATCATAAACTGTTTAACTCTTTTTTTTCC	2239
Db	48041	CTGTAAACATTTGACTCTGTATTTTCAACAAATCATAAACTGTTTAACTCTTTTTTTTCC	48100
Qy	2240	AAAAAAAAAAAAA 2252	

```
Db 48101 AACCTGGAAGAA 48113

RESULT 10
AY166504
LOCUS
DEFINITION
Brassica rapa subsp. pekinensis Myb-related transcription factor
ACCESSION
AY166504
VERSION
AY166504.1 GI:27357049
SOURCE
Brassica rapa subsp. pekinensis
ORGANISM
Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
FEATURES
1. (bases 1 to 318)
REFERENCE
Jang,H. and Hur,Y.
AUTHORS
Study on the expression pattern of genes associated with flowering
TITLE
in chinese cabbage plants
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 318)
AUTHORS
Jang,H. and Hur,Y.
TITLE
Direct Submission
JOURNAL
Submitted (16-Oct-2002) Biology, Chungnam National University,
Kungdong, Yuseong-ku, Deajeon 305-764, Korea
LOCATION/Qualifiers
1. 318
/organism="Brassica rapa subsp. pekinensis"
/mol_type="mRNA"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/tissue_type="leaf"
<1..>318
/codon_start=1
/product="Myb-related transcription factor CCA1"
/protein_id="AAN86551.1"
/db_xref="GI:27357050"
/translation="GQPTQYPMHIMPLVPLGSSITRSLSHAPDENSEHSHPHTVAGDHQS
FPDHIMSSLQTPALYTAASFASFSTWPPDTSVGPQNNLTPLAAMASATVAASAAW
AANG"
BASE COUNT 78 a 93 c 67 g 80 t
ORIGIN
Query Match 9.9%; Score 223.4; DB 8; Length 318;
Best Local Similarity 83.2%; Pred. No. 4e-42;
Matches 267; Conservative 0; Mismatches 51; Indels 3; Gaps 1;
QY 895 CAAGGACCTCAGACTTATCCGATGCATATCCCTGTGCTAGTGCCTATGGGAGCTCAATA 954
Db 1 CAAGGACCTCAGACTTATCCGATGCATATCCCTGTGCTAGTGCCTATGGGAGCTCAATA 60
QY 955 ACAAGTCTCTATCATCCTCTTCAGAGCCAGATAGTATCCCTCCACACAGTTGCGAGA 1014
Db 61 ACAAGTCTCTGTACATGCTCCAGACGAGATAGTATCATCCCTCCACACAGTTGCGAGC 120
QY 1015 GATTATCAGTCGTTTCTTAATCATATAATGTCAACCTTTTACAAACACCGGCTCTTTAT 1074
Db 121 GATCATCAGTCGTTTCCGACCATATATGTCAAGCTTTTACAGACACAGCTCTATAC 180
QY 1075 ACTGGCGCACTTTTGGCTCATCATATTTGGCTCTCCGATTCCTAGTGGTGGCTCACTGTT 1134
Db 181 ACAGCTGCAAGTTTGGCTTCAACGTTTGGCTCTCTGCTACTAGTGTGGCCCAACAG-- 238
QY 1135 CCAGGGAACCTACCTCCGAATCTGGCTGCGATGGCGGAGCCACTGTTGCACTGTAGT 1194
Db 239 -GGAAACAATTAACTCCGAATCTGGCTGCGATGGCTTTCAGCCACTGTTGCACTGCAAGT 297
QY 1195 GCTTGGTGGGCTGCCAATGGA 1215
Db 298 GCTTGGTGGGCTGCCAATGGA 318

RESULT 11
A68428
LOCUS
DEFINITION
Sequence 1 from Patent WO9749811.
ACCESSION
A68428
VERSION
A68428.1 GI:4759505
SOURCE
unidentified
ORGANISM
unidentified
REFERENCE
1 (bases 1 to 2526)
AUTHORS
Coupland,G.M. and Schaffer,R.J.
TITLE
GENETIC CONTROL OF FLOWERING
JOURNAL
Patent: WO 9749811-A 1 31-DEC-1997;
INNES JOHN CENTRE INNOV LTD (GB)
FEATURES
Location/Qualifiers
1..2526
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
338..2275
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAB42406.1"
/db_xref="GI:4759506"
/translaton="MDNTSGEELLAARKPKYTIITKORERWTEDEHERFLEALRLYGR
AWQRIEHHIGTKTAVQIRSHAQKFFTKLEKEAEVKGIPVCOALDIEIPPRPKRPNT
PIPRKPGNGTSSQVSSAKDAKLVSSASSQNLQAPLDLKERMPFSEKSTGTGKENDQ
NCSGVSTVNYKPLPTKQVSGDIETSKTVDNAVQVPKKNKDKDNGDGTTHSMQNY
PWHFADI VNGNI AKCPQNPQSMVQDFMFHMPREETHGHANLQTASGATTTASHQ
APFACHQDDYRSFLQISSTFNLINLQNPAAHAATFAASVAPYASVGNSSGDS
TPMSSSPSITAIATAATAATAAWASHGLLPVCAPAPITCVPSFSTVAVPTBAMTEND
TVENTQPFQKNTALQDQTLASKSPASSDDSDSETGVTKLNADSKTNDKIEEVVTA
AVHDSNTAKNLVDKSSCSNTSPSSDAETDALDKMKDKEDVKTENDKQPDVIELN
NRKIKMRDNNNNNATTDWSKEVSEEGRIAFQALFARERLPQSPFPPQVAENVRKOS
DTSMPLAPNFKSQDSCAADQGVVMIGVGTCKSLKTRQTGTFKPKYKRCSEMEVKEQVGN
INQSDKVKCKRLRLEGEAST"
BASE COUNT 767 a 523 c 528 g 708 t
ORIGIN
Query Match 9.6%; Score 217.2; DB 6; Length 2526;
Best Local Similarity 49.9%; Pred. No. 1e-40;
Matches 986; Conservative 0; Mismatches 823; Indels 165; Gaps 11;
QY 235 GTGATGGACACAAATTCGCTCTGGAGAGATCTGGTTATTAAAGACTCGGAAGCCATATACG 294
Db 335 GTTATGGATGACTAATAACATCTCGAGAGAAATATTATTAGCTAAGGCAAGAAAGCCATATACA 394
QY 295 ATAACAAAGCAACGTTGAAAGGTGGACTGAGGAGAAACATAATAGATTTCATTGAAGCTTTG 354
Db 395 ATAACAAAGCAGGAGAGCGATGGACTGAGGATGAGCATGAGAGGTTCTTAGAAGCCTTG 454
QY 355 AGGCTTTATGGTAGACATGGCAGAGATTGAAGAACATGTAGCAACAAACAACTGCTGTC 414
Db 455 AGGCTTTATGGAAGAGCTTTGGCAACGAAATTGAAGAACATATTGGGACAAAGACTGCTGTT 514
QY 415 CAGATAAGAAGTCAGCTCAGAAAATTTTCTCAAGGTAGAGAAAGAGGCTGAAGCTAAA 474
Db 515 CAGATCAGAAGTCATGCACAAAAGTTCTTCACAAAAGTTGGAGAAAGAGGCTGAAGTTAAA 574
QY 475 GGTGTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCCTTAAGCGTAAA 534
Db 575 GGCATCCCTGTTTGGCAAGCTTTGGACATAGAAAATTCGCGCTCTCTCGTCTCAACGAAA 634
QY 535 CCAAAACAATCTTATCTCTCGAAAGACGGGAAGTGAACGATCCCTTATGTCAAAAACCGGT 594
Db 635 CCAATATCTCTTATCTCTCGAAAACCTGGGAACACCGGTACATCTTCTCTCAAGTATCA 694
QY 595 GTGAATGATGGAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGATGGCAAT 654
Db 695 TCAGCAAAAGATGCAAAAACCTTGTTCATCGGCTCTCTTCTTCAAGTGAATCAGGCGTTC 754
QY 655 GAAGATCGACAAACAATCAAGCCTGAAGAGAAAC-----TCTGCAGGAA 699
```



```
Db      755  TTGGATTGGAATAAATCGCGTTCTCTGAGAAACATCAACTCGAAAGAAAATCAAGAT 814
QY      700  GACAACTGTTGAGATTGTTTCACTCATCATGATCTCTGCTGCTCCATCGAATAAA 759
Db      815  GAGAATTGCTCGGGTGTCTTCTACTGTGAACAAGTATCCCTTACCAACGAAACAGGTAAGT 874
QY      760  AGTTGTATAGAGATCAAAACGCAAGCATTCTT-----CCGCGAGTTCTTGCTTC 809
Db      875  GCGGACATTGAAACAAGTAGACCTCAACTGTGGAACAACGCGGTTCAAGATGTTCCAAAG 934
QY      810  ACGGGAAGAGGGAGTCAAGATTAACAGGGTA-----AGAAAGGAGTCAAAAC 855
Db      935  AAGAAACAAAGACAAAGATGTTAAGATGTTACTGTCACAGCATGCAAAAACACTACCCCT 994
QY      856  TCAGATTGGAATCGAAAATCTCTGGAATAACGGTAATGAGCAAGGACCTCAGACTTATCC- 914
Db      995  TGGCATTTCCAGCAGATATTGTGAAGGGAATATAGCAAAATGCCCCCAAAATCATCCC 1054
QY      915  -----G 915
Db      1055  TCAGGTATGGTATCTCAAGACTTTCATGTTTTCATCCTTAGAGAGAAGAACTACGGGCAC 1114
QY      916  ATGCATATCCCTGTGCTAGTGCCATTGGGAGCTCAATTAACAAGTTCTCTATCACAATCCT 975
Db      1115  GCAAAATCTTCAAGCTACAACAGCATCTGCTACTACTACAGCTTCTCATCAAGCGTTTCCA 1174
QY      976  CTTTCAGAGCCAGATAGTCAATCCCAACACAGATGTTGCGAGGAGATTATCAGTCGTTTCTAAT 1035
Db      1175  GCTTGTCAATCAGAGATGATACCGTTGTTTCTCCAGATATCATCTTCTTCCAAAT 1234
QY      1036  CATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTATATCGCCGCAACTTTTCGGCTCA 1095
Db      1235  CTTATTATGTCAACTCTCTACAGAATCTCTGCAGCTCATGCTGCAGCTACATTTCGCTGCT 1294
QY      1096  TCAATTTGGCTCCCG-----ATTCTAGTGGTGGCTCACCTGTTCCAGGGAAC 1143
Db      1295  TCGGCTGCGCTTATCGAGTGTGCGGAATTTCTGGTATTCTATCAACCCCAATGAGCTCT 1354
QY      1144  TCACCTCCGAATCTGGCTGCGATGCGCGAGCACCTGTTGCAGCTCTAGTGTCTTGGTGG 1203
Db      1355  TCTCTCCAGTATAACTGCCATTGCGCTGCTACAGTAGTCTGCTCAACTGTTGGTGG 1414
QY      1204  GTCGCAATGGATTATTAACCTTTATGTGCTCTCTTATGTTCAAGTGGTGTTCATAGTCAT 1263
Db      1415  GCTTCTCATGGACTTCTTCTGTATGCGTCCAGCTCCAAATACATGTGTTCATCTTCA 1474
QY      1264  CTTCCATCTACTTTGGACCATCATGTGATGATGATGATGATGATGATGATGATGATGATGAT 1323
Db      1475  ACTGTTGCAATCCAACTCCAGCAATGACTGAAATGGATACCGTTGAAAATATCTCAACCG 1534
QY      1324  CATGGTTCTGTGAGAGCCGAGAGCAAGAACACTCCGAGGCATCAAAAGGCTCGATCTTCA 1383
Db      1535  TTTGAGAAACAAACACAGCTCTGCAAGATCAAACTTGGCTTCAAAATCTCAGCTTCA 1594
QY      1384  CTGAGCTCAGAGATGTTGAAAT-----AAGAGTAAACCACTTTGTCTATGAGCAG 1434
Db      1595  TCATCTGTGATGATTCAGATGAGCTGGAGTAACCAAGCTAAATGCCGACTCAAAAACCAAT 1654
QY      1435  CTTCTGCAACACCTCAGAGTATGCA-----AGGTTTCAGATGAGGAGCAGGA 1482
Db      1655  GATGATAAAATTTAGAGAGGTTGTTGTTATGTCGCGCTGTGATGATGATGATGATGATGAT 1714
QY      1483  GACAGAAAACAAAGTTGACCGGCTCTCGTGTGGCTCAAAACACTCCGTCGAGTAGTATGATGAT 1542
Db      1715  AAGAAAAATCTTGTGACCGCTCATGCTGTGGCTCAAAATACACCTTCAGGAGTAGCAGA 1774
QY      1543  GTTGAGCGGATGATCAAGAAAGGCAAGAGGATGGCAACCAATGGTGTGATGAAAGAAACG 1602
Db      1775  GAAACTGATGATTTAGATAAAATGGAGAAAGATTAAGAGGATGTGAAGGAGACAGATGAG 1834
QY      1603  AAT-----GAAGACACTAATAAACCCTCAAACTTCAGAGTCCAAATGCAAGCCGCGCAGTAGA 1656
```

```
Db      1835  AATCAGCCAGATGTTATTGAGTTAAATTAACCGTAAGATTAAATGAGAGACAACAACAGC 1894
QY      1657  ATCAGCTCCAATATAATACCGATCCATGGAAGTCTGTGCTGACGAGGTCGAATTCCTTC 1716
Db      1895  AACAAACAATGCAACTACTGATTCGTGGAAGGAAGTCTCCGAAGAGGGTCGTATAGCGTTT 1954
QY      1717  CAAGCTCTCTCTTCCAGAGAGGATTATTCGCCCAAGTTTACATATCGAGAAGAACACAGA 1776
Db      1955  CAGGCTCTCTTTCAGAGAGAAGATTGCTTCAAGCTTTTCGGCTCTCTCAAGTGGCAGAG 2014
QY      1777  GAGGAAGAACAACAACAACAAGATATCCAAATGCGACTTGTATCTTAACTTCAACA 1836
Db      2015  AATGTGAATAGAAAAACAAGTGCACGT-----CAATGCCATTGGCTCTCTAATTTCAA 2068
QY      1837  GCTCAGTTAACACCACTGATGATCAAGAGGAGAGAAACACAGAGATTCTTGGATC 1896
Db      2069  AGCCAGGATCTTGTGCTGCGACCAAGA-----AGGAGTAGTAATGATC 2113
QY      1897  GGATTAGATGCTTCAAGCTAATGAGTAGAGGAAGAACAGGTTTAAACCCATACAAAGA 1956
Db      2114  GGTGTTGGAACATGCAAGAGTCTTAAACGAGACACAGAGATTTAAGCCATACAGAGA 2173
QY      1957  TGTTCATGGAAGCCAAAGAGTAGAATCTCTCAACAACAATCTTATCATTTCTGTGAA 2016
Db      2174  TGTTCATGGAAGTGAAGAGAGAGCCAAAGTTGGGAACATAAA-----CAATCAAAGTGT 2227
QY      2017  CAGAAAGATCCCAACCGGATGCGTTGGAACTCAAGCTTCCACATGAGACTCT 2070
Db      2228  GAAAAGTCTGCAAAAGGCTTCGATTGGAAGGAGAGCTTCTACATGACAGACT 2281
```

```
RESULT 12
ATAJ6404
LOCUS   Arabidopsis thaliana mRNA for LATE ELONGATED HYPOCOTYL MYB
DEFINITION
ACCESSION AJ006404.1 GI:3281845
VERSION   Arabidopsis thaliana (thale cress)
KEYWORDS  transcription factor.
SOURCE    Arabidopsis thaliana
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS   Schaffer,R., Ramsay,N., Samach,A., Corden,S., Putterill,J.,
          Carre,I.A. and Coupland,G.
          The late elongated hypocotyl mutation of Arabidopsis disrupts
          circadian rhythms and the photoperiodic control of flowering
          Cell 93 (7), 1219-1229 (1998)
JOURNAL   98319236
MEDLINE   9657154
PUBMED    2 (bases 1 to 2526)
REFERENCE
AUTHORS   Samach,A.
          Direct Submission
          Submitted (26-MAY-1998) Samach A., Molecular Genetics, John Innes
          Centre, Colney Lane, Norwich NR4 7UH, UK
JOURNAL
FEATURES
source    1. .2526
          /organism="Arabidopsis thaliana"
          /mol_type="mRNA"
          /db_xref="taxon:3702"
          1..2526
          /gene="late elongated hypocotyl"
          /gene="late elongated hypocotyl"
          /function="MYB transcription factor"
          /codon_start=1
          /protein_id="CAA07004.1"
          /db_xref="GI:3281846"
          /db_xref="SPTREMBL:O81713"
          /translation="MDTNTSGEELLAARKRPYITTKQERWTFDEHERLEALRLYGR
          AQRIEIHGKTAVQIRSHAQKFFTKLEAEVKGIPVCQALDIIEIPPRPKRNT
```

PYPRKPGNNGTSSQVSSAKDAKLVSASSQLNQAFLDLKMPFSEBKTSTGKENDQDE  
NCSGVSTVNYKPLPTKQVSGDIETSTVDNAVDQVPKKNKDKDNDGTVHSHOMVY  
PMFHADIVNGNIAKQFNHPSGMVSDQFMPHMRSETHGHANLQATYASATTTASHQ  
APFACHSODDYFSFLOJSSFTSNLIMSTLLONPAHAATAFAASVMPYASVNGSDSS  
TPWSSSPSITAJAAATVAATAWASHGILLPVCAPIATTCVPFSTVAVPNTAMD  
TVENTQFBEKONTALQOQTLASKSPASSDDSDDETGVTKLNADSKNDKIBEVVITA  
AVHDSNTAQKNLVDKSSCGSNTSGSDAETDADLDMKEDKEDKEDENQDPVIELN  
NRKTKMDNNSNNATTDSSWKEVEEGRIAFQALFARERLPQSPFPQVAENVNRKOS  
DTSMLAPNFKSQSDCAADQEGVVMIGVTGCKSLKTRQTFKPKYKRCMEVKESQVGN  
INNOSEKVKCKRLRLEGEAST"

BASE COUNT 767 a 523 c 528 g 708 t  
ORIGIN

Query Match 9.6%; Score 217.2; DB 8; Length 2526;  
Best Local Similarity 49.9%; Pred. No. 1e-40;  
Matches 986; Conservative 0; Mismatches 823; Indels 165; Gaps 11;

QY	235	GTGATGAGACAAATTCGTCTGGAGAAGATCGTGTATTAGACTCGGAAGCCATATACG	294
DB	335	GTATGGATACTAATACATCTCGAGAAGAATTATTAGCTAAGGAAGAAAGCCATATACA	394
QY	295	ATAAAGCAACCGTGAAGGTGACTGAGGAAGACATATAGATTCAATTTGAAGCTTG	354
DB	395	ATAAAGCAGCAGAGCGATGACTGAGGATGAGCATGAGAGTTTCTAGAAGCCTTG	454
QY	355	AGGCTTTATGGTAGAGCATGCGAAGATTGAAGAACATGAGCAACAAAACCTGCTGTC	414
DB	455	AGGCTTTATGGAGAGCTTGGCAACGAATTCGAAGAACATATTGGGCAAGAGACTGCTGT	514
QY	415	CAGATAGAAGTACCGTCAAGATTTTCTCCAGGTAGAGAAGAGGCTGAAGCTAAA	474
DB	515	CAGATAGAAGTATGACAAAAGTTCTTCAAAAGTTGAGAGAAGAGGCTGAAGTTAAA	574
QY	475	GGTGTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACCGCCTTAAGCGTAAA	534
DB	575	GGCATCCTGTTTGCCHAGCTTTGGACATAGAAATTCGGCTCTCTGCTTAAACGAAA	634
QY	535	CCAAACAACTCTTATCTCGAAGACGGGAAGTGGAAACGATCTCTTATGTCAAAAACGGT	594
DB	635	CCCAATACTCTTATCTCGAAGACCTTGGGAACAAACGGTACATCTTCTCAAGTATCA	694
QY	595	GTGAATCATGGAAGAGTCCCTTGGATCAGAAAAGTGTGCGATCTGAGATGGCCAA	654
DB	695	TCAGCAAAAGATGCAAAACTGTTTTCATCGGCTCTTCTTCAAGTTTGAATCAGGCGTTC	754
QY	655	GAAGATCGACAACAACTCAAGCGCTGAAGAGAAAAC-----TCTGCAGGAA	699
DB	755	TTGGATTGGAATAAATGCGCTTCTCTGAGAAAACATCAACTGGAAAAGAAAATCAAGAT	814
QY	700	GACAACTGTTTCAAGTGTTCATCTCATCATGATATCTCTCTGCTGCATCTCTCCATGAATA	759
DB	815	GAGAAATGTCTGGGTGTCTTACTGTGAACAAGTATCCCTTACCAACGAAACAGGTAAGT	874
QY	760	AGTTGTATAGACATCAAGCGAAGACATTT-----CGGCGAGTTCTTGCCTTC	809
DB	875	GGCGACATTTGAACAAGTAAGACTCAACTGTGGACAAACCGGTTCAGAGATGTTCCCAAG	934
QY	810	ACGGGAAGAGGGAAGTCAAGATTAACAGGGTA-----AGAAAAGGAGTCAAAAC	855
DB	935	AAGAACAAAGACAAAGATGTAACGATGTTACTTCTGTCACAGCATGCAAAACTACCTT	994
QY	856	TCAGATTTGATGCAAAATCTCTGGAAAACGGTAATGAGCAAGGACCTCAGACTTATCC-	914
DB	995	TGGCATTTTCCACGACAGATATTGTGAACGGGAATATAGCAAAATGCGCTCAAAATCATCC	1054
QY	915	-----G	915
DB	1055	TCAGGTATGGTATCAAGACTTCATGTTTCTATCTCTATGAGAGAAGAACTCAGGGCAC	1114
QY	916	ATGCATATCCCTGTGTAGTGGCAATTTGGGAGCTCAATAACAAAGTTCTCTTATCACATCCT	975
DB	1115	GCAAACTCTTCAAGCTACAAAGCATCTGCTTACTTACTTACAGCTTCTCATCAAGGTTTCCA	1174

QY	976	CTTTCAGAGCCAGATAGTCAATCCACACAGTTGCGAGAGATTATCAGTCGTTTTCTCTAAT	1035
DB	1175	GCTTGTCAATCACAGATGATTACCGTTGTTTTCTCCAGATATCATCTACTTTCTCCCAAT	1234
QY	1036	CATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTATACTGCGCGCAACTTTCGCCTCA	1095
DB	1235	CTTATTATGTCAACTCTCTCTACAGAACTCTGCGAGCTCATGCTGCAGCTACATTCGCTGCT	1294
QY	1096	TCATTTTGGCTCTCCG-----ATTCTAGTGGTGGCTACCTGTTCCAGGGAAAC	1143
DB	1295	TCGCTTGGCCCTTATGCGAGTGTGCGGAATTTCTGGTGATTATCAACACCCCAATAGCTCT	1354
QY	1144	TCACCTCGGAATCTGGCTGCCATGGCGCAGCCACTGTTGCGAGCTGCTAGTCTTGGTGG	1203
DB	1355	TCTCTCCAAGTATAACTGCGCATTTGCCGCTCTACAGTAGCTGCTGCAACTCTTGGTGG	1414
QY	1204	GCTGCCAATGGATTATTAACCTTTTATGCTCTCTTCTTATTGTTTCTAGTGGTGGTTCAC	1263
DB	1415	GCTTCTCATGAGCTTCTTCTGTTATGCTCCAGTCCCAATTAACATGTTTCCATCTCTCA	1474
QY	1264	CCTCATCTACTTTTGGACCATCATGTGTAGAGTACACAAAGCAAGCACTTTTACAA	1323
DB	1475	ACTGTTGCAGTTCOAATCTCCAGCAATGACTGAAATGGATACCGTTTGAATAATACTCAAC	1534
QY	1324	CATGTTTCTGTGACAGCCGAGAGCAAGAACACTCCGAGGCGATCAAAAGGCTCGATCTTCA	1383
DB	1535	TTTGAGAAAACAAACACAGCTCTGCAAGATCAAACTTGGCTTTCOAATCTCCAGCTTCA	1594
QY	1384	CTGACTCAGAGGATGTTTGAATAAT-----AAGAGTAAACCAAGTTTGTCTAGCAGCAG	1434
DB	1595	TCATCTGATGATTGAGATGAGCTGGAGTAACCAAGCTAAATGCCACTCAAAACCAAT	1654
QY	1435	CTTCTCGAACCTTGAGAGTGTGCAA-----AGGGTTTCAAGTGGAGCAGGA	1482
DB	1655	GATGATAAAATTTGAGGAGGTTGTTGTTACTGCGCTGTGTGATGACTCAAAACACTGCC	1714
QY	1483	GACAGAAAACAAAGTTGACCGTCTCGTGTGGCTCAACACTCCGTCGAGTAGTGTATGAT	1542
DB	1715	AAGAAAATCTTGTGGACCGCTCATGCTGTGGCTCAATACACTTTCAGGAGGTGACCGA	1774
QY	1543	GTTGAGCGGATGCATCAGAAAGCAAGAGATGGCCAAATGTTGAGGTGAAAGAAACG	1602
DB	1775	GAAACTGATGATTTAGATAAAATGGAGAAAGATAAGAGGATGTGAAGGACAGATGAG	1834
QY	1603	AAT-----CAAGACACTATAAACCTCAAACTTCAGAGTCCAATGCACGCCGAGTAGA	1656
DB	1835	AATCAGCCAGATGTTATTGAGTTAAATTAACCGTAAAGATTAAATTCAGAGACAACACAGC	1894
QY	1657	ATCAGCTTCCAAATATAACCGATCCATGGAAGTCTGTGTCGACGAGGTCGAATTTGCCCTTC	1716
DB	1895	AACAACATGCAACTACTGATTCGTGAAGAAAGTCTCCGAAGAGGTCGTATAGCGTTT	1954
QY	1717	CAAGCTCTCTTCTCAGAGAGGTTATTGCCCGCAAGTTTTACATATCGAAGAAACACAGA	1776
DB	1955	CAGGCTCTCTTTCAGAGAAAGATTGCTTCAAGCTTTTTCGCTCTCTCAAGTGGCAGAG	2014
QY	1777	GAGGAAGAACAAACAAACAAAGAAATATCCAAATGGCACTTTGATCTTAACTTCACTACA	1836
DB	2015	AATGTGAATAGAAAACAAAGTGACCGT-----CAATGCCATTTGGCTCTCTAAATTTCAA	2068
QY	1837	GCTCAGTTAACACCACTGTTGATGATCAAGAGGAGAGAAACACAGGATTTCTTGGATC	1896
DB	2069	AGCCAGGATCTTGTGCTGACAGACCAAGA-----AGGAGTAGTAATGATC	2113
QY	1897	GGATTAGATGCTTCAAGCTTAATGAGTAGAGGAGAAACAGGTTTTTAAACCATACAAAAGA	1956
DB	2114	GGTGTGGAACATGCAAGAGTCTTAAACCGAGACAGACAGAGATTTAAGCCATACAGAGA	2173
QY	1957	TGTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAACTCTTATCTATCTTGTGGAA	2016
DB	2174	TGTTCAATGGAAGTGAAGAGAGCCAAAGTTGGGAACATAAA-----CAATCAAGTGTAT	2227
QY	2017	CAGAAAGATCCCAACCGGATCGGTTTGGAAACTCAAGCTTCCACATGAGACTCT	2070

```
Db      2228  GAAAAAGTCTGCRAAAGGCTTCGATTGGAAGGAGAGCTTCTACATCAGAGACT 2281
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
RESULT 13
AX506421 1938 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 1116 from Patent WO0216655.
DEFINITION AX506421
ACCESSION AX506421
VERSION AX506421.1 GI:23387658
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE Harper, J.F., Kreps, J., Wang, X., and Zhu, T.
AUTHORS Stress-regulated genes of plants, transgenic plants containing
TITLE same, and methods of use
JOURNAL Patent: WO 0216655-A 1116 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source
1.1938
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 632 a 417 c 418 g 471 t
ORIGIN
Query Match 9.4%; Score 213; DB 6; Length 1938;
Best Local Similarity 49.4%; Pred. No. 1e-39;
Matches 980; Conservative 0; Mismatches 820; Indels 165; Gaps 11;
Qy 238 ATGAGACAAATTCGTCTGGAGAAGATCTGGTTATTAAAGCTCGGAAGCCATATACGATA 297
Db 1 ATGATATACTATACATCTCGGAAGAAGATTATTAGCTAAGGCAAGAAAGCCATATACAA 60
Qy 298 ACAAGCAACTGAAAGTGACATGAGCAAGCAATATAGATTCAATGAAGCTTTCAGG 357
Db 61 ACAAGCAGCAGAGCAGATGACATGAGGATGAGCATGAGAGTTTCTAGAAAGCTTTCAGG 120
Qy 358 CTTTATGGTATGAGATGGCAAGATTGGAAGACATGTAGCAACAAATTCGTCTCCAG 417
Db 121 CTTTATGGAAGAGCTTGGCAACGAATTGAAGAACATATTGGCAACAGAGCTGCTGTTCA 180
Qy 418 ATAAGAAGTCAGCTCAGAAATTTTCTCAGAGTAGAAGAGGCTGAAGCTAAAGGT 477
Db 181 ATCAGAAGTCATGCACAAAAGTTCTTCACAAAGTTGGAGAAAGAGGCTGAAGTTTAAAGGC 240
Qy 478 GTAGCTATGGTCAAGGCTAGACATAGCTATTCTCTCCACGGCCTAAGCGTAAACCA 537
Db 241 ATCCCTGTTGGCAAGCTTTGGACATAGAAATTCGGCTCTCTGCTTAACGAAACCC 300
Qy 538 AACAAATCCTTATCTCGAAAGACGGGAAGTGGAAACGATCTCTTATGTCAAAAACGGGTGTG 597
Db 301 AATACTCCTTATCTCGAAAGCCTGGGAACAAACGGTATCATCTTCTCTCAAGTATCATCA 360
Qy 598 AATGATGGAAGAGTCCCTTGGATCAGAAAAGATGTCGATCTCGATGGCCCAATGAA 657
Db 361 GCAAAAGATGCAAAACTTTGTTTCATCGGCCCTCTTCTTCACAGTTGAAATCAGGCGTCTTG 420
Qy 658 GATCGACAACAAATCAAGCCTGGAAGAGAAAC-----TCTCGAGGAAGAC 702
Db 421 GATTGGAAAAAATGCGGCTTCTTGAGAAAACATCACTGGAAAAGAAAATCAAGATGAG 480
Qy 703 AACTGTTTCAGATGTTTTCACATCATCAGTATCTCTCTGTCGATCCTCCATGAATTAAGT 762
Db 481 AATTGCTCGGGTGTCTTCTACTGTGAACAAAGTATCCCTTACCAACGAAACAGGTAAGTGGC 540
Qy 763 TGATATAGACATCAACCGCAAGACATTTC-----GCCAGTTCTTGCCCTTCAG 812
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

Qy	1780	GAAGAA	CAACAA	CAACAA	CAAGAA	CAAGAG	ATATCC	CAATGG	CACCTT	GATCTT	TAACCT	CACAGCT	1839		
Db	1681	GTGNA	TAGAAA	CAAA	GTGAC	CGT-----	CAATGC	CAAT	TGCTCT	TAATTT	CAAAAGC	1734			
Qy	1840	CAGTTA	CACCGT	TGATG	ATCA	AGGAG	GAGAA	CACAG	GAATTT	CTTGG	AAATCGGA	1899			
Db	1735	CAGAT	TTCTTG	CTGC	GAGAC	CAAGA-----	AGAG	TAGT	AAATG	ATGAT	CGGT	1779			
Qy	1900	TTAGAT	GTCTT	CAAA	GGCTA	ATGAG	TAGAG	GAGAA	CAGGTTT	TAACCAT	CAAAAAGATG	1959			
Db	1780	GTTT	GGAAT	CTG	CAAG	AGCTTT	AAAA	CGAG	ACAG	CAGGATTT	TAAGCC	ATCAAGAGATG	1839		
Qy	1960	TCCAT	CGAG	CCCA	AGAA	AGT	TAGAT	CTCT	CAAG	CAAG	CAATCCT	TATCAT	CTTGGAA	CAG	2019
Db	1840	TCAAT	GGAGT	GAA	AGAG	CC	CAAGTTT	GGGAA	CAATAAA-----	CAAT	CAAA	GTGATGAA	1893		
Qy	2020	AAAGAT	CCCCAA	CGGAT	TCGG	TTGG	AAACT	CAAG	CTTCC	CATGA	2064				
Db	1894	AAAGT	CTGCA	AAAG	GGCTT	CCGAT	TTGGA	AGGAG	GAAGCTT	CTCAT	CA	1938			

RESULT 14	PVU420902	2172 bp mRNA linear	PLN 23-MAY-2002
LOCUS	Phaseolus vulgaris mRNA for LHY protein.		
DEFINITION	Phaseolus vulgaris mRNA for LHY protein.		
ACCESSION	AJ420902		
VERSION	AJ420902.2 GI:21213867		
KEYWORDS	LHY protein.		
SOURCE	Phaseolus vulgaris		
ORGANISM	Phaseolus vulgaris		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.		
REFERENCE	1		
AUTHORS	Kaldis, A.		
TITLE	Reciprocal control between the light-induced acute response and the following circadian cycle results in the entrainment of the clock in Phaseolus vulgaris		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2172)		
AUTHORS	Prombona, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-NOV-2001) Prombona A., Institute of Biology, Nat.Center for Scient. Res.'Demokritos', Terma Patriarkou Grigoriou, 15310 Ag. Paraskevi, Attiki, GREECE		
REMARK	revised by author [20-MAY-2002]		
FEATURES	Location/Qualifiers		
	1..2172		

```

/organism="Phaseolus vulgaris"
/mol_type="mRNA"
/variety="Red Kidney"
/db_xref="taxon:3885"
/tissue_type="leaves"
/dev_stage="seedlings"
1..2172
/gene="lhy"
1..2172
/gene="lhy"
/function="putative transcription factor"
/codon_start=1
/product="LHY protein"
/protein_id="CAD12767.2"
/db_xref="GI:21213868"
/translational="MDAVSSGEVVVTKRKYITIKQRRWTEEHNRFLKALVKLHGR
AQRTEEHGCTVAQIRSHAQKFFTKLEKALVKGVPJQCALDIDIPPRPKRKPSPN
PFRKTIITATLHSGAKGNLVESHNNQALDKEPEKPYDIDEGLTITVKNKDE
NCSGVFKVQEPVPCSSISSANRSTSSMSPVGLNSCVLKEITSSVKEVITRDENTSFL
TVELGNRLNEIDNGQANGTSKNSTLENSDALQTKLQVNEKTDGGLDSALTIQMGQNG
NYPRVYVTHVVDGKLGTSQNPQSQDMLFRDSMFQPIGGDNGQPNLFTNSAPNTSESQ
NNTARSSVHQSELPYPPPTQHNQDDYQSFELHSSSTNLNVSTLLQNPAAHVAASFAA
TWPVYANPETSADSPRCSQGGTSRQISQPPSVAATAATVAAAATAAAHAGGLPLCL

```

BASE COUNT	704 a	466 c	481 g	521 t	
ORIGIN					
Query Match	9.0%;	Score 202.2;	DB 8;	Length 2172;	
Best Local Similarity	75.2%;	Pred. No. 3.8e-37;			
Matches 252;	Conservative	0;	Mismatches 83;	Indels 0;	Gaps 0;
Qy	238	ATGGAGACAAATTCGTTCTCGGAAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA	297		
Db	1	ATGGACGCATCTCCTCTCGGAAAGAAAGTCGTTGTTAAACAGCAGAAAAACCGTATACGATC	60		
Qy	298	ACAAAGCAACGTTGAAGGTGGACTGAGGAAGAACATAATAGATTCTATTGAAAGCTTTTGAGG	357		
Db	61	ACAAAGCAAGGGAACGATGGACAGAGGAGGAGCATATAGGTTCTTAGAAGCTTTTGA	120		
Qy	358	CTTTATTGGTAGAGCATGGCAGAAAGATTGGAAGAACATGTAGCAACAAAAAATCTGCTCCAG	417		
Db	121	CTGCACGGCGAGCATGGCAGCGCATAGAAGAGCATATAGGAACAAAGACTCCCGTGCAA	180		
Qy	418	ATAAGAAGTCAAGCTCAGAAATTTTCTCCAAGGTAGAGAAAGAGCGCTGAAGCTAAAGGT	477		
Db	181	ATCAGAGAGTCAGCAGACAAGTTCTTTACAAAGTTGGAGAAAGAGGCGCCCTTGTAAAGG	240		
Qy	478	GTAGCTATGGGTCAAGCGCTAGACATAGCTATTCTCTCTCCACGGCTAAGCGTAAACCA	537		
Db	241	GTTCCAATTGGACAAGCTCTTGACATAGATATTCCTCCACGGCCCAAGAAACCA	300		
Qy	538	AACAATCCTTATCTCGAAGACGGGAAGTGGAAAC	572		
Db	301	AGCAATTCCTTATCTCGAAGACCGAGTTGGTAC	335		

RESULT 15					
AX509497					
LOCUS	AX509497	1134 bp	DNA		
DEFINITION	Sequence 4192 from Patent WO0216655.				
ACCESSION	AX509497				
VERSION	AX509497.1	GI:23390734			
			linear	PAT 27-SEP-2002	

```

KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.
1
REFERENCE
AUTHORS
Harpser,J.F., Kress,J., Wang,X. and Zhu,T.
TITLE
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL
Patent: WO 0216655-A. 4192 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source
Location/Qualifiers
1. .1134
/mol_type="genomic_DNA"
/db_xref="taxon:3702"
BASE COUNT 333 a 193 c 214 g 394 t
ORIGIN
Query Match 8.4%; Score 189; DB 6; Length 1134;
Best Local Similarity 100.0%; Pred.No. 5.5e-34;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGAGATTTCCTCCATTTCGGTAGCTCTCGTCTCTTTCTTTGTTTCATTGATCAAAAGCA 60
Db 814 TGAGATTTCCTCCATTTCGGTAGCTCTCGTCTCTTTCTTTGTTTCATTGATCAAAAGCA 873
QY 61 AATCACTTCTTCTTCTTCTTCTTCGCGATTCTTACTGTTTCTTATCCAAAGAAATCTG 120

```

Search completed: November 24, 2003, 09:24:33  
Job time : 8068.27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2003, 02:51:36 ; Search time 1266.07 Seconds  
(without alignments)  
11223.496 Million cell updates/sec

Title: US-10-084-553-1

Perfect score: 4344  
Sequence: 1 gcagtggttcacttaaga.....actcataagtaagcacaaaa 4344

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 2172232 seqs, 163554964 residues

Total number of hits satisfying chosen parameters: 434464

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4344	100.0	4344	14	US-10-084-553-1
2	1527.6	35.2	2254	14	US-10-084-553-3
3	1525.2	35.1	2240	10	US-09-887-576-617
4	1349.2	31.1	1827	10	US-09-938-842A-1500
5	1137	26.2	1137	10	US-09-887-576-702
6	1134	26.1	1134	10	US-09-938-842A-4192
7	312.6	7.2	865	9	US-09-770-445-578
8	289	6.7	479	11	US-09-770-961-961
9	83.2	1.9	2526	11	US-09-934-455-119
10	83.2	1.9	2526	15	US-10-278-536-199
11	81.6	1.9	1938	10	US-09-938-842A-1116
12	66.4	1.5	3673778	12	US-10-312-841-1
13	65.4	1.5	8524	12	US-10-311-455-1814
14	62.4	1.4	15373	12	US-10-311-455-439
15	61.6	1.4	525	10	US-09-938-842A-2341
16	59.8	1.4	17594	12	US-10-311-455-1999

17	59.6	1.4	12138	12	US-10-311-455-1915	Sequence 1915, Ap
18	59.6	1.4	12138	12	US-10-240-453-209	Sequence 209, App
19	59.4	1.4	6115	12	US-10-311-455-1774	Sequence 1774, Ap
20	58.4	1.3	411	10	US-09-960-352-14521	Sequence 14521, A
21	57.4	1.3	6095	12	US-10-311-455-334	Sequence 334, App
22	57.4	1.3	6095	12	US-10-240-485-28	Sequence 28, Appl
23	57	1.3	7455	12	US-10-311-455-1732	Sequence 1732, Ap
24	56.8	1.3	113515	12	US-10-311-455-500	Sequence 500, App
25	56.6	1.3	113515	12	US-10-311-455-2148	Sequence 2148, Ap
26	56.4	1.3	1164	11	US-09-934-455-247	Sequence 247, App
27	56.4	1.3	8303	12	US-10-311-455-395	Sequence 395, App
28	56.4	1.3	3673778	12	US-10-312-841-2	Sequence 2, Appli
29	56	1.3	13326	12	US-10-311-455-1585	Sequence 1685, Ap
30	55.8	1.3	413	10	US-09-960-352-2919	Sequence 2919, Ap
31	55.6	1.3	5610	12	US-10-204-708-54	Sequence 54, Appl
32	55.6	1.3	5610	12	US-10-240-453-132	Sequence 192, App
33	55.6	1.3	5610	14	US-10-239-676-170	Sequence 170, App
34	55.4	1.3	6125	12	US-10-311-455-1586	Sequence 1586, Ap
35	55.4	1.3	6125	12	US-10-240-453-152	Sequence 152, App
36	55.4	1.3	113515	12	US-10-311-455-2147	Sequence 2147, Ap
37	55.2	1.3	7304	12	US-10-204-708-44	Sequence 44, Appl
38	55.2	1.3	40862	12	US-10-311-455-2045	Sequence 2045, Ap
39	54.8	1.3	17183	12	US-10-311-455-459	Sequence 459, App
40	54.6	1.3	7047	12	US-10-240-453-260	Sequence 260, App
41	54.6	1.3	7346	12	US-10-311-455-318	Sequence 318, App
42	54	1.2	5763	12	US-10-311-455-1261	Sequence 1261, Ap
43	54	1.2	6191	12	US-10-311-455-1190	Sequence 1190, Ap
44	53.6	1.2	15387	12	US-10-311-455-157	Sequence 157, App
45	53.4	1.2	6352	12	US-10-311-455-878	Sequence 878, App

#### ALIGNMENTS

#### RESULT 1

US-10-084-553-1  
; Sequence 1, Application US/10084553  
; Publication No. US20030056247A1  
; GENERAL INFORMATION:  
; APPLICANT: Tobin, Elaine  
; APPLICANT: Wang, Zhi-Yong  
; APPLICANT: Sun, Lin  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Phytochrome Regulated Transcription Factor for Control  
; TITLE OF INVENTION: of Higher Plant Development  
; FILE REFERENCE: 023070-124200US  
; CURRENT APPLICATION NUMBER: US/10/084,553  
; CURRENT FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: US 08/843,572  
; PRIOR FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4344  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: phytochrome-regulated transcriptional factor CCA1  
; OTHER INFORMATION: genomic clone  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1332)..(1370)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1371)..(1448)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1449)..(1560)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1561)..(1648)  
; NAME/KEY: exon

```

; LOCATION: (1649)..(1710)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1711)..(2189)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2190)..(2384)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (2385)..(2471)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2472)..(2645)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (2646)..(2728)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2729)..(3610)
; FEATURE:
; NAME/KEY: intron
; LOCATION: (3611)..(3680)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (3681)..(4043)
; US-10-084-553-1

Query Match      100.0%; Score 4344; DB 14; Length 4344;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGTGGTTCACITACAAAGAACCTGGTCTTCAAAACAGACAGAGGTAAACCAATTTCTCTCTT 60
DB 1 GCAGTGGTTCACITACAAAGAACCTGGTCTTCAAAACAGACAGAGGTAAACCAATTTCTCTCTT 60

QY 61 TAACCTGTGTTTGGTGGTGAATTAATCTAGAGTGAAGAACTCAAAATTCAGAGAAAT 120
DB 61 TAACCTGTGTTTGGTGGTGAATTAATCTAGAGTGAAGAACTCAAAATTCAGAGAAAT 120

QY 121 TGTGTTGTTATCTGTTTCAGGAGGCTTTGTTTCAGAGAGGTCAAGACACATACAAAGACA 180
DB 121 TGTGTTGTTATCTGTTTCAGGAGGCTTTGTTTCAGAGAGGTCAAGACACATACAAAGACA 180

QY 181 TATTAGGAGCAGCTGAATCAAAAGGAGAAAGAAAGAAAGAGGCTTTTTCAGAGCCAT 240
DB 181 TATTAGGAGCAGCTGAATCAAAAGGAGAAAGAAAGAAAGAGGCTTTTTCAGAGCCAT 240

QY 241 TCATGAATTCGAATGAAGGATATCAAAAGAACTCAAAAGGCCACGCTTCCTTCA 300
DB 241 TCATGAATTCGAATGAAGGATATCAAAAGAACTCAAAAGGCCACGCTTCCTTCA 300

QY 301 ATCTTTCCCTTCTGTAACATAAATTTTCATCCTTTCTCTCTCTCTCTCTCTCTCTCTT 360
DB 301 ATCTTTCCCTTCTGTAACATAAATTTTCATCCTTTCTCTCTCTCTCTCTCTCTCTCTT 360

QY 361 TTTAGCTCAAAGTATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGACTATATAT 420
DB 361 TTTAGCTCAAAGTATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGACTATATAT 420

QY 421 GAGATGTTTGTTCATTTTCCAAAATTTCAAACCTTTGCCCATTTAGTCTTACCCT 480
DB 421 GAGATGTTTGTTCATTTTCCAAAATTTCAAACCTTTGCCCATTTAGTCTTACCCT 480

QY 481 TCATGATGGTTAGCTTATAGCTGAACCTGTTGAATAACGATATGGCCCTTATGCT 540
DB 481 TCATGATGGTTAGCTTATAGCTGAACCTGTTGAATAACGATATGGCCCTTATGCT 540

QY 541 AAAAGAACAAACCTTATGGGCTTAAAAAATAAGCCCAATATAAACTATATGGCCCAA 600
DB 541 AAAAGAACAAACCTTATGGGCTTAAAAAATAAGCCCAATATAAACTATATGGCCCAA 600

QY 601 TAAAGTTAGGTCATAGGTGTGAGATAGCGGTGTAGTGAACCGCACGAGATGCGC 660
DB 601 TAAAGTTAGGTCATAGGTGTGAGATAGCGGTGTAGTGAACCGCACGAGATGCGC 660

661 GTTCGATTGTTGGTGAAGTAGTCGTCTAGATTCCCGGTCCACTGATGTTTCTAGTGTAT 720
661 GTTCGATTGTTGGTGAAGTAGTCGTCTAGATTCCCGGTCCACTGATGTTTCTAGTGTAT 720
721 CAGACACGTCGTCACAAAACCTGGTGGGAGAGATTAAACGATCTTAAGTAGGTCCCACTAGAT 780
721 CAGACACGTCGTCACAAAACCTGGTGGGAGAGATTAAACGATCTTAAGTAGGTCCCACTAGAT 780
781 CAAGATATTAATAACGAATTGACCTTTTAAACCTTTTCAAGGTAGTCCCGAACTCGTGGCCT 840
781 CAAGATATTAATAACGAATTGACCTTTTAAACCTTTTCAAGGTAGTCCCGAACTCGTGGCCT 840
841 AGAATACAAAGAAGGTTTGCAACAAGTTGATGTTAGATGACACAGAAATGTAACCTGAAC 900
841 AGAATACAAAGAAGGTTTGCAACAAGTTGATGTTAGATGACACAGAAATGTAACCTGAAC 900
901 AAAAGCTGAATCATCTCTTCAGCCACTAGTATGTTGACATATGGCAGTTCCTTTTGTAGC 960
901 AAAAGCTGAATCATCTCTTCAGCCACTAGTATGTTGACATATGGCAGTTCCTTTTGTAGC 960
961 CTCGAAATAAATTAATTAAGGTTTTCAGGTTTAAAGATAATTAATGAGTGGCTGAGATTTCT 1020
961 CTCGAAATAAATTAATTAAGGTTTTCAGGTTTAAAGATAATTAATGAGTGGCTGAGATTTCT 1020
1021 CCATTTCCGTAGCTTCTGCTCTCTTTTCTTTTCTTTTCAATGATCAAAAGCAAACTCCTTCT 1080
1021 CCATTTCCGTAGCTTCTGCTCTCTTTTCTTTTCTTTTCAATGATCAAAAGCAAACTCCTTCT 1080
1081 TCTTCTTCTTCTCTCGATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1140
1081 TCTTCTTCTTCTCTCGATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1140
1141 TCGAATCTTTATCGAATCCAAGCTGATTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
1141 TCGAATCTTTATCGAATCCAAGCTGATTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
1201 TACTTAAGATTGATTTATTTGTGTCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1260
1201 TACTTAAGATTGATTTATTTGTGTCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1260
1261 TTGTTTTTTTGTGTTGGATTAGTGGAAATTTCTGTAAGAGAAAGATCTGAGTTCTGTGAGA 1320
1261 TTGTTTTTTTGTGTTGGATTAGTGGAAATTTCTGTAAGAGAAAGATCTGAGTTCTGTGAGA 1320
1321 GGAGCTTAGTGTAGGACAAATTCGTCCTGAGAGAAATCTGGTTATTTAAGGTAAATTAAC 1380
1321 GGAGCTTAGTGTAGGACAAATTCGTCCTGAGAGAAATCTGGTTATTTAAGGTAAATTAAC 1380
1381 TAAATTTTAGGGGGAAGATGATTTGTTTAGGTCTCAAGATTCGAGAAATTTTAAATGAAACT 1440
1381 TAAATTTTAGGGGGAAGATGATTTGTTTAGGTCTCAAGATTCGAGAAATTTTAAATGAAACT 1440
1441 TGATATAGACTCGGAAGCCATATACGATAACAAAGCAACGTAAGAGTGGCTGAGGAG 1500
1441 TGATATAGACTCGGAAGCCATATACGATAACAAAGCAACGTAAGAGTGGCTGAGGAG 1500
1501 AACATAATAGATTCATTGAAGCTTTGAGGCTTTATGTTAGAGCATGGCAGAAAGATTCAG 1560
1501 AACATAATAGATTCATTGAAGCTTTGAGGCTTTATGTTAGAGCATGGCAGAAAGATTCAG 1560
1561 GTTGATTTTATTTCCCTTTATATGCTTATTTTGTGTTTTCAGAGGTTTGTCTTCAA 1620
1561 GTTGATTTTATTTCCCTTTATATGCTTATTTTGTGTTTTCAGAGGTTTGTCTTCAA 1620
1621 ACTGATTTGCTTTTTTTTCAATTTTGGACAGAACTATGAGCAACAAACTGCTGTCACAGATA 1680
1621 ACTGATTTGCTTTTTTTTCAATTTTGGACAGAACTATGAGCAACAAACTGCTGTCACAGATA 1680
1681 AGAAGTCACGCTCAGAAATTTTCTCCTCAAGGTAATTCGGTTAATTTGAAATGATGTTTC 1740
1681 AGAAGTCACGCTCAGAAATTTTCTCCTCAAGGTAATTCGGTTAATTTGAAATGATGTTTC 1740
```



QY	1741	TCATCTTCATTTGGCTTAATGCTTAAGACTTAATGAAAGCCAGGCCAAGTTTTTCGCTTCCTT	1800
DB	1741	TCATCTTCATTTGGCTTAATGCTTAAGACTTAATGAAAGCCAGGCCAAGTTTTTCGCTTCCTT	1800
QY	1801	TTGCTTCTTTAGTCAGGAGATAGATAGATTACGTTTTTAGAGCTTTAGTAATGACGAATAAG	1860
DB	1801	TTGCTTCTTTAGTCAGGAGATAGATAGATTACGTTTTTAGAGCTTTAGTAATGACGAATAAG	1860
QY	1861	TCTTAAAAATGTTGGAGAAATGACGAGATGAATCGTTTTCTTTGTTTTATGCCATAATC	1920
DB	1861	TCTTAAAAATGTTGGAGAAATGACGAGATGAATCGTTTTCTTTGTTTTATGCCATAATC	1920
QY	1921	TTGTTAAATCCACAACATGTACATAGATCTTTCAGAGAAATGTTAGTTTTCTTTAGATTCT	1980
DB	1921	TTGTTAAATCCACAACATGTACATAGATCTTTCAGAGAAATGTTAGTTTTCTTTAGATTCT	1980
QY	1981	TCAGATAAACTTTGTGTCCTTTCTTACCGATTCTGAGGTAGTGCAAAAAGTGCGCTGAGTGCT	2040
DB	1981	TCAGATAAACTTTGTGTCCTTTCTTACCGATTCTGAGGTAGTGCAAAAAGTGCGCTGAGTGCT	2040
QY	2041	AGAAATTTTTGAAATGTTCCCTTGTGTGATTAAGCCATAGAGGTAAACCATTTTTTCCAG	2100
DB	2041	AGAAATTTTTGAAATGTTCCCTTGTGTGATTAAGCCATAGAGGTAAACCATTTTTTCCAG	2100
QY	2101	TTCTGTCAATTTAAACTTGTAGTGCTCATTAGATTTTTTGTGTTTACGTTTGTTTAGAG	2160
DB	2101	TTCTGTCAATTTAAACTTGTAGTGCTCATTAGATTTTTTGTGTTTACGTTTGTTTAGAG	2160
QY	2161	GGTAAACAAACTACTCTCATCTCTCAGGTAGAGAAAGCGCTGAAAGCTTAAAGGTGTAG	2220
DB	2161	GGTAAACAAACTACTCTCATCTCTCAGGTAGAGAAAGCGCTGAAAGGTGTAG	2220
QY	2221	CTATGGGTCAAGCGGTAGACATAGCTATTCTCTCTCCAGCGCTAAAGCGTAAACCAACA	2280
DB	2221	CTATGGGTCAAGCGGTAGACATAGCTATTCTCTCTCCAGCGCTAAAGCGTAAACCAACA	2280
QY	2281	ATCCTTATCTCTCGAAAGACGGGAAGTGGACAGTCTTTATGTCAAAAACGGGTGTGAATG	2340
DB	2281	ATCCTTATCTCTCGAAAGACGGGAAGTGGACAGTCTTTATGTCAAAAACGGGTGTGAATG	2340
QY	2341	ATGAAAAGAGTCCCTTTGGATCAGAAAAAGTGTGCGCATCTCTGAGGTGATTTTTCATGGTCA	2400
DB	2341	ATGAAAAGAGTCCCTTTGGATCAGAAAAAGTGTGCGCATCTCTGAGGTGATTTTTCATGGTCA	2400
QY	2401	TATGGCATCTTTTTTGCAGTGTGCACATTTGCTCTCTCATGTTTATTAATACAGATTGTGTGC	2460
DB	2401	TATGGCATCTTTTTTGCAGTGTGCACATTTGCTCTCTCATGTTTATTAATACAGATTGTGTGC	2460
QY	2461	TTCTGTTTATAGATGCGCAATCAAGATCGACAACAATCAAAAGCTTGAAGAGAAACTCTGC	2520
DB	2461	TTCTGTTTATAGATGCGCAATGAAAGATCGACAACAATCAAAAGCTTGAAGAGAAACTCTGC	2520
QY	2521	AGGAAGACAACCTGTTTCAGATTGTTTCTACTCATCAGTATCTCTCTGCTGCACTCTCCATGA	2580
DB	2521	AGGAAGACAACCTGTTTCAGATTGTTTCTACTCATCAGTATCTCTCTGCTGCACTCTCCATGA	2580
QY	2581	ATAAAAGTTGTATAGAGACATCAAAACGAGCACTTTCCGCGAGTTCTTGCTTTCACGGG	2640
DB	2581	ATAAAAGTTGTATAGAGACATCAAAACGAGCACTTTCCGCGAGTTCTTGCTTTCACGGG	2640
QY	2641	AAGAGGTAAAAAACAATCTTTTCAATTGCTATTGAGGTTTTTAAGACGATTAGTACTTTTCA	2700
DB	2641	AAGAGGTAAAAAACAATCTTTTCAATTGCTATTGAGGTTTTTAAGACGATTAGTACTTTTCA	2700
QY	2701	TGAAACTAAAAACCGTGGGGGAATAACAGGGAAGTCAGAAATAACAGGGTAAAGAAAGGATC	2760
DB	2701	TGAAACTAAAAACCGTGGGGGAATAACAGGGAAGTCAGAAATAACAGGGTAAAGAAAGGATC	2760
QY	2761	AAACTCAGATTTGAAATGCAAAATCTCTGGAAAAACGGTAAATGACGAAGCACTCAGACTTA	2820
DB	2761	AAACTCAGATTTGAAATGCAAAATCTCTGGAAAAACGGTAAATGACGAAGCACTCAGACTTA	2820

QY	2821	TCCGATGCATATCCCTGTGCTAGTGGCCATTGGGGAGCTCAATAAACAGTTCTCTATCACA	2880
Lb	2821	TCCGATGCATATCCCTGTGCTAGTGGCCATTGGGGAGCTCAATAAACAGTTCTCTATCACA	2880
QY	2881	TCCTCCTTCAGAGCCAGATAGTCATCCCAACACAGTTGTCAGGAGATTATCAGTCGTTTC	2940
Db	2881	TCCTCCTTCAGAGCCAGATAGTCATCCCAACACAGTTGTCAGGAGATTATCAGTCGTTTC	2940
QY	2941	TAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTATCTGCGCCAACTTCGC	3000
Db	2941	TAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTATCTGCGCCAACTTCGC	3000
QY	3001	CTCATCATTTTGGGCTCCGATTCAGTGTGGTCACTGTTTCAGGGAACTCACTCTC	3060
Db	3001	CTCATCATTTTGGGCTCCGATTCAGTGTGGTCACTGTTTCAGGGAACTCACTCTC	3060
LY	3061	GAATCTGGCTGCATGCGCCGAGCCACTGTTGAGAGCTGCTAGTGGCTTGGTGGGCTGCCAA	3120
Db	3061	GAATCTGGCTGCATGCGCCGAGCCACTGTTGAGAGCTGCTAGTGGCTTGGTGGGCTGCCAA	3120
QY	3121	TGGATTATTACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGTCATCTCCATC	3180
Db	3121	TGGATTATTACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGTCATCTCCATC	3180
QY	3181	TACTTTTGGACCATCATGTGATGTAGAGTACACAAAGCAAGCACATTACAAATGGTTT	3240
Db	3181	TACTTTTGGACCATCATGTGATGTAGAGTACACAAAGCAAGCACATTACAAATGGTTT	3240
QY	3241	TGTCAGAGCCGAGAGCAAGAACACTCCGAGGCATCAAAAGCTCGATCTTCACTGACATC	3300
Db	3241	TGTCAGAGCCGAGAGCAAGAACACTCCGAGGCATCAAAAGCTCGATCTTCACTGACATC	3300
QY	3301	AGAGGATGTTGAAATAAGAGTAAACACAGTTTGTCTATGAGCAGCCCTTCTGCAACACTGA	3360
Db	3301	AGAGGATGTTGAAATAAGAGTAAACACAGTTTGTCTATGAGCAGCCCTTCTGCAACACTGA	3360
QY	3361	GAGTGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAAACAAAGTTGACCGGTCCTCGTG	3420
Db	3361	GAGTGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAAACAAAGTTGACCGGTCCTCGTG	3420
QY	3421	TGGCTCAAAACACTCCGTGAGTGTGATGATGTTGAGGCGGATGCAATCAGAAAGGCAAGA	3480
Db	3421	TGGCTCAAAACACTCCGTGAGTGTGATGATGTTGAGGCGGATGCAATCAGAAAGGCAAGA	3480
QY	3481	GGATGGCCACCAATGGTGCAGGTGAAAGAAACGAAATGAAGACACTAAATAAACCTTCAACTTC	3540
Db	3481	GGATGGCCACCAATGGTGCAGGTGAAAGAAACGAAATGAAGACACTAAATAAACCTTCAACTTC	3540
QY	3541	AGAGTCCAATGCAACCGCCAGTGAATCAGCTCCAATATAACCGATCCATGGAAGTCTGT	3600
Db	3541	AGAGTCCAATGCAACCGCCAGTGAATCAGCTCCAATATAACCGATCCATGGAAGTCTGT	3600
QY	3601	GTCTGACGAGGTACTTACTTGGACTAAAGATCAACTTCCTTTATTTTCAATCATTTTCTC	3660
Db	3601	GTCTGACGAGGTACTTACTTGGACTAAAGATCAACTTCCTTTATTTTCAATCATTTTCTC	3660
QY	3661	ATATAAATATTGTACATTTCGGGTGCAATTGCTCTTCCAAGCTCTCTTCTCCAGAGAGTAT	3720
Db	3661	ATATAAATATTGTACATTTCGGGTGCAATTGCTCTTCCAAGCTCTCTTCTCCAGAGAGTAT	3720
QY	3721	TGCCGCAAAAGTTTTACATATCGAAGAAACACAGAGAGGAAGAACAAACAAACAGAAC	3780
Db	3721	TGCCGCAAAAGTTTTACATATCGAAGAAACACAGAGAGGAAGAACAAACAAACAGAAC	3780
QY	3781	AAAGATATCCAAATGGCACTTGATCTTAACTTCACAGCTCAGTTTAAACACAGTTTGATGATC	3840
Db	3781	AAAGATATCCAAATGGCACTTGATCTTAACTTCACAGCTCAGTTTAAACACAGTTTGATGATC	3840
QY	3841	AAGAGGAAGAAGAAACACAGGATTTCTTGGAAATCGATTAGATGCTTCAAAGCTTAATGA	3900
Db	3841	AAGAGGAAGAAGAAACACAGGATTTCTTGGAAATCGATTAGATGCTTCAAAGCTTAATGA	3900
QY	3901	GTAGAGGAAGAACAGGTTTTTAAACCAATACAAAGATGTTCCATGGAAGCCAAAGAAAGTA	3960



Db 1471 GATGGAGCAGGAGACAGAAAAAAGTTGACCGGTCTCTGTGGCTCAAAACACTCCGTGC 1530  
QY  
3440 AGTAGTGATGATGTTGAGGCGGATGCATCAGAAAGCGAAGAGATGCGCAACCAATGGTGAG 3499  
Db 1531 AGTAGTGATGATGTTGAGGCGGATGCATCAGAAAGCGAAGAGATGCGCAACCAATGGTGAG 1590  
QY 3500 GTGAAAGAAACGAATGAAGACACTAATAAACCTCAACCTTCAAGAGTCCAATGACACGCCG 3559  
Db 1591 GTGAAAGAAACGAATGAAGACACTAATAAACCTCAACCTTCAAGAGTCCAATGACACGCCG 1650  
QY 3560 AGTAGAATCAGCTCCCAATATAACCGATCCATGAGTCTGTGCTGACGAGGTACTTACT 3619  
Db 1651 AGTAGAATCAGCTCCCAATATAACCGATCCATGAGTCTGTGCTGACGAGGTACTTACT 1700  
QY 3620 TGGACTAAAGATCAACTTCTTTATTTCAAAATCAATTTCTCATATAAAATATTGTACATTC 3679  
Db 1701 ----- 1700  
QY 3680 GGTGCAATGCTTCCCAAGCTCTTCTCCAGAGAGGTATTCGCGCAAAAGTTTTACATA 3739  
Db 1701 GGTGCAATGCTTCCCAAGCTCTTCTCCAGAGAGGTATTCGCGCAAAAGTTTTACATA 1760  
QY 3740 TCCAGAGAACACAGAGAGGAGAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 3799  
Db 1761 TCCAGAGAACACAGAGAGGAGAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1820  
QY 3800 TGAATCTTAACCTTCCACAGCTCAGTTAAACCCAGTTGATGATCAAGAGAGAGAGAGAGAAC 3859  
Db 1821 TGAATCTTAACCTTCCACAGCTCAGTTAAACCCAGTTGATGATCAAGAGAGAGAGAGAAC 1880  
QY 3860 AGGATTTCTTGGAAATCGGATAGATGCTTCAAGCTTAATGAGTAGAGAGAGAGAGAGTT 3919  
Db 1881 AGGATTTCTTGGAAATCGGATAGATGCTTCAAGCTTAATGAGTAGAGAGAGAGAGAGTT 1940  
QY 3920 TAAACCATACAAAGATGTTCCATGGAAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3979  
Db 1941 TAAACCATACAAAGATGTTCCATGGAAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2000  
QY 3980 TATCATTTGATGGAACAGAAAGATCCCAACCGGATGCGGTTGGAAACTCAAGCTTCCAC 4039  
Db 2001 TATCATTTGATGGAACAGAAAGATCCCAACCGGATGCGGTTGGAAACTCAAGCTTCCAC 2060  
QY 4040 ATGAGACTATATTTTCATCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4099  
Db 2061 ATGAGACTATATTTTCATCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2120  
QY 4100 CTGCTACATTTCTTTTCTTTTGGGCGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4159  
Db 2121 CTGCTACATTTCTTTTCTTTTGGGCGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2180  
QY 4160 TGTAACTTTGACTCTGTATTATTTCAACAAATCAATAAAGTGTGTTGTTGTTGTTGTTGTT 4219  
Db 2181 TGTAACTTTGACTCTGTATTATTTCAACAAATCAATAAAGTGTGTTGTTGTTGTTGTTGTT 2240  
QY 4220 ACCTGGAAGAA 4231  
Db 2241 AAAAAAAAAAAAA 2252

RESULT 3

US-09-887-576-617  
; Sequence 617, Application US/09887576  
; Patent No. US20020144047A1  
; GENERAL INFORMATION:  
; APPLICANT: Budworth, P.  
; APPLICANT: Brown, D.  
; APPLICANT: Chang, H.  
; APPLICANT: Zhu, T.  
; APPLICANT: Han, B.  
; APPLICANT: Wang, X.  
; APPLICANT: Cooper, Bret  
; TITLE OF INVENTION: Promoters for regulation of plant expression  
; FILE REFERENCE: 1360.001US1

; CURRENT APPLICATION NUMBER: US/09/887,576  
; CURRENT FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: US 60/213,848  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/214,087  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/258,692  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 875  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 617  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; US-09-887-576-617

Query Match 35.1%; Score 1525.2; DB 10; Length 2240;  
Best Local Similarity 88.1%; Pred. No. 0;  
Matches 1797; Conservative 0; Mismatches 3; Indels 240; Gaps 3;

QY 2180 TCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTTAAGGTGTAGCTATGGTCAAGCGCTAGA 2239  
Db 441 TTTCTCAAGGTAGAGAAAGAGGCTGAAGCTTAAGGTGTAGCTATGGTCAAGCGCTAGA 500  
QY 2240 CATAGCTATTCTCTCCACGGCTTAAGCGTAAACCAACCAATCCTTATCTCCTCAAGAC 2299  
Db 501 CATAGCTATTCTCTCCACGGCTTAAGCGTAAACCAACCAATCCTTATCTCCTCAAGAC 560  
QY 2300 GGAAGTGAACGATCCTTATGTCAAAAACGGGTGAAATGATGGAAGAGTCCCTTGG 2359  
Db 561 GGAAGTGAACGATCCTTATGTCAAAAACGGGTGAAATGATGGAAGAGTCCCTTGG 620  
QY 2360 ATCAGAAAAGTCTCCATCCTCAGGTGATTTTCATGGTCTATATGGCATCTTTTTCAGT 2419  
Db 621 ATCAGAAAAGTCTCCATCCTCAGGTGATTTTCATGGTCTATATGGCATCTTTTTCAGT 643  
QY 2420 GTGTCACTTGTCTCTCATGTTATTAATACAGATTTGTGCTGTTTATAGATGGCCAA 2479  
Db 644 -----AGATGGCCAA 653  
QY 2480 TGAAGATCGACAACAATCAAGCTGAAGAGAAAACCTCTCAGAGAAACAACTGTTCAGA 2539  
Db 654 TGAAGATCGACAACAATCAAGCTGAAGAGAAAACCTCTCAGAGAAACAACTGTTCAGA 713  
QY 2540 TTGTTTCACTCATCATGATCTCTCTGCTGATCCTCCATGAATAAAGTTGTATAGAGAC 2599  
Db 714 TTGTTTCACTCATCATGATCTCTCTGCTGATCCTCCATGAATAAAGTTGTATAGAGAC 773  
QY 2600 ATCAAAACGGAAGCACATTTCCGCGAGTTCTTCCCTTCAACGGGAAGAGGTAACAAATCT 2659  
Db 774 ATCAAAACGGAAGCACATTTCCGCGAGTTCTTCCCTTCAACGGGAAG----- 817  
QY 2660 TTCAITGCTATTGAGGTTTTAAGACGATTTAGTACTTTTTCATGAACACTAAAAACCGTGGG 2719  
Db 818 ----- 817  
QY 2720 GAATAACAGGAGTCTAGAAATACAGGTTAAGAGAGTCAAACTCAGATTTGAATGCA 2779  
Db 818 -----AGGGAAGTCTAGAAATACAGGTTAAGAGAGTCAAACTCAGATTTGAATGCA 870  
QY 2780 AAATCTCTGGAACCGGTAAATGAGCAAGGACCTCAGACTTATCCGATGATATCCCTGTG 2839  
Db 871 AAATCTCTGGAACCGGTAAATGAGCAAGGACCTCAGACTTATCCGATGATATCCCTGTG 930  
QY 2840 CTAGTCCCATTTGGGAGCTCAATAACAAGTTCTCTATCATATCCTCTCCTCAGAGCCAGAT 2899  
Db 931 CTAGTCCCATTTGGGAGCTCAATAACAAGTTCTCTATCATATCCTCTCCTCAGAGCCAGAT 990  
QY 2900 AGTCATCCCCACACAGTTGCGAGAGATTATCAGTCTTTCCTTAATCATATATGTCAACC 2959  
Db 991 AGTCATCCCCACACAGTTGCGAGAGATTATCAGTCTTTCCTTAATCATATATGTCAACC 1050  
QY 2960 CTTTTCACAAACACCGGCTCTTTTACTGCGGCAACTTTTCGCGCTCATCATTTTGGCGCTCC 3019

Db	2061	ATGAGACTCTATTTTTCATCTGATCTGTGTGTTGTGACTGTGTTTTTAAAGTTTTCAGAGACCA	2122
Qy	4100	CTGCTACATTTTCTTTTTTCTTTTGGAGCCCTTTGTATTGTTTCCCTGTGTCATAGTCTTCC	4159
Db	2121	CTGCTACATTTTCTTTTTTCTTTTGGAGCCCTTTGTATTGTTTCTGTGTCATAGTCTTCC	2180
Qy	4160	TGTAACAATTTGACTCTGTATTATTTCACAAATCATAAACTGTTTAAATCTTTTTTTTCCA	4219
Db	2181	TGTAACAATTTGACTCTGTATTATTTCACAAATCATAAACTGTTTAAATCTTTTTTTTCCA	2240
RESULT 4			
US-09-938-842A-1500			
; Sequence 1500, Application US/09938842A			
; Patent No. US20020160378A1			
; GENERAL INFORMATION:			
; APPLICANT: Harper, Jeff			
; APPLICANT: Kreps, Joel			
; APPLICANT: Wang, Xun			
; APPLICANT: Zhu, Tong			
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONT			
; TITLE OF INVENTION: SAME, AND METHODS OF USE			
; FILE REFERENCE: SRIPI300-3			
; CURRENT APPLICATION NUMBER: US/09/938,842A			
; CURRENT FILING DATE: 2001-08-24			
; PRIOR APPLICATION NUMBER: US 60/227,866			
; PRIOR FILING DATE: 2000-08-24			
; PRIOR APPLICATION NUMBER: US 60/264,647			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/300,111			
; PRIOR FILING DATE: 2001-06-22			
; NUMBER OF SEQ ID NOS: 5379			
; SEQ ID NO 1500			
; LENGTH: 1827			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
US-09-938-842A-1500			
Query Match			
Best Local Similarity 31.1%; Score 1349, 2; DB 10; Length 1827;			
Matches 1621; Conservative 0; Mismatches 3; Indels 240; Gaps 3;			
Qy	2180	TCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGTAGCTATGGGTCAAGCGCTAGA	2239
Db	204	TTTCTCAAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGTAGCTATGGGTCAAGCGCTAGA	263
Qy	2240	CATAGCTATTCTCTCCACGGCTAAGCGTAAACCAACAATCTTATCTCTCGAAGAC	2299
Db	264	CATAGCTATTCTCTCCACGGCTAAGCGTAAACCAACAATCTTATCTCTCGAAGAC	323
Qy	2300	GGGAAGTGGAAACCATCTTATGTCAAAAACGGGTGTGAATGATGGAAGAGTCCCTTGG	2359
Db	324	GGGAAGTGGAAACCATCTTATGTCAAAAACGGGTGTGAATGATGGAAGAGTCCCTTGG	383
Qy	2360	ATCAGAAAAAGTGTGCGATCCTCAGGTGATTTTTCCTGATCATGTCATCTTTTTCAGT	2419
Db	384	ATCAGAAAAAGTGTGCGATCCTG-----406	
Qy	2420	GTGTACATGTCTCTCATGTATTATTAACAGATTGTGTCTTCTGTTATAGATGCCAA	2479
Db	407	-----AGATGGCCAA 416	
Qy	2480	TGAAGATCGAACAAATCAAGCCTGAAGAGAAAACTCTGCAGGAAGACAACCTGTTTCTAGA	2539
Db	417	TGAAGATCGAACAAATCAAGCCTGAAGAGAAAACTCTGCAGGAAGACAACCTGTTTCTAGA	476
Qy	2540	TTGTTTCACTCATCAGTATCTCTCTGTCATCCTCCATGAATAAAGTTGTATAGAGAC	2599
Db	477	TTGTTTCACTCATCAGTATCTCTCTGTCATCCTCCATGAATAAAGTTGTATAGAGAC	536
Qy	2600	ATCAAAACGCAAGCACTTTCCGCGAGTCTTTGCTCTCAGCGGAAGAGTAAAAAACAATCT	2659
Db	537	ATCAAAACGCAAGCACTTTCCGCGAGTCTTTGCTCTCAGCGGAAGAGTAAAAAACAATCT	580

```
QY 2660 TTCAATGCTATTGAGGTTTAAAGACGATAGTAGTCTTTTCATGAAACTAAAAACCGTGGG 2719
Db 581 ----- 580

QY 2720 GAATAACAGGAAGTCAGATAACAGGGTAAGAAAGGAGTCAAACTCAGATTGGAATGCA 2779
Db 581 ----- AGGAAAGTCAGAAATAACAGGGTAAGAAAGGAGTCAAACTCAGATTGGAATGCA 633

QY 2780 AAATCTCTGAAAAACGGTAATGAGCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG 2839
Db 634 AAATCTCTGAAAAACGGTAATGAGCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG 693

QY 2840 CTAGTGCCATTGGGAGCTCAATAACAAGTTCTCTATCACATCCTCTCTCAGAGCCAGAT 2899
Db 694 CTAGTGCCATTGGGAGCTCAATAACAAGTTCTCTATCACATCCTCTCTCAGAGCCAGAT 753

QY 2900 AGTCATCCCCACACAGTTGCAGGAGATTATCAGTCGTTTCCCTAATCATATAATGTCAACC 2959
Db 754 AGTCATCCCCACACAGTTGCAGGAGATTATCAGTCGTTTCCCTAATCATATAATGTCAACC 813

QY 2960 CTTTACAAACACCGGCTCTTTATCTGCGCAACTTTTCGCCCTCATCTTTTGGCCTCCC 3019
Db 814 CTTTACAAACACCGGCTCTTTATCTGCGCAACTTTTCGCCCTCATCTTTTGGCCTCCC 873

QY 3020 GATTCTAGTGGTGCCTCACCTGTTCCAGGAACTCACTCCGAATCTGGCTGCCATGGCC 3079
Db 874 GATTCTAGTGGTGCCTCACCTGTTCCAGGAACTCACTCCGAATCTGGCTGCCATGGCC 933

QY 3080 GCAGCCACTGTTGCAGCTGCTAGTGTCTGTTGGTGGCTGCCAATGGATTATACCTTTATGT 3139
Db 934 GCAGCCACTGTTGCAGCTGCTAGTGTCTGTTGGTGGCTGCCAATGGATTATACCTTTATGT 993

QY 3140 GCTCCTCTTAGTTTCAAGTGGTGTTCACCTAGTCACTCTCCATCTACTTTTGGACCATCATGT 3199
Db 994 GCTCCTCTTAGTTTCAAGTGGTGTTCACCTAGTCACTCTCCATCTACTTTTGGACCATCATGT 1053

QY 3200 GATGTAGATACAAAAAGCAAGCATTTTACAACATGGTTCTGTGAGAGCCGAGAGCAA 3259
Db 1054 GATGTAGATACAAAAAGCAAGCATTTTACAACATGGTTCTGTGAGAGCCGAGAGCAA 1113

QY 3260 GAACACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAATAAG 3319
Db 1114 GAACACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAATAAG 1173

QY 3320 AGTAAACCAAGTTTGTATGAGCAGCCCTTCTGCAACACCTGAGAGTGATGCAAAAGGTTCA 3379
Db 1174 AGTAAACCAAGTTTGTATGAGCAGCCCTTCTGCAACACCTGAGAGTGATGCAAAAGGTTCA 1233

QY 3380 GATGAGCAGGAGACAGAAAAAAGTTGACCGGTCTCTGTGTGGCTCAAAACACTCCGTCG 3439
Db 1234 GATGAGCAGGAGACAGAAAAAAGTTGACCGGTCTCTGTGTGGCTCAAAACACTCCGTCG 1293

QY 3440 AGTAGTGATGATGTTGAGGCGGATGATCAGAAAGGCAAGAGATGCGCAATGTGTAG 3499
Db 1294 AGTAGTGATGATGTTGAGGCGGATGATCAGAAAGGCAAGAGATGCGCAATGTGTAG 1353

QY 3500 GTGAAAGAAACGAATGAAGACACTAAATAAACCTTCAAACTTCAAGTCCCAATGACGCCCG 3559
Db 1354 GTGAAAGAAACGAATGAAGACACTAAATAAACCTTCAAACTTCAAGTCCCAATGACGCCCG 1413

QY 3560 AGTAGAATCAGTCCCAATATACCGATCCATGGAAGTCTGTGTCTGACGAGGCTACTTACT 3619
Db 1414 AGTAGAATCAGTCCCAATATACCGATCCATGGAAGTCTGTGTCTGACGAGGCTACTTACT 1463

QY 3620 TGGACTTAAAGATCAACTTCTTTTAAATCATTTTCTCATATAAATTTGTACATTC 3679
Db 1464 ----- 1463

QY 3680 GGGTCGAATTTGCCCTTCCAAAGCTCTCTTCTCCAGAGAGGTATTGCCGCAAGTTTACATA 3739
Db 1464 GGGTCGAATTTGCCCTTCCAAAGCTCTCTTCTCCAGAGAGGTATTGCCGCAAGTTTACATA 1523
```

```
QY 3740 TCAGAGAAGAACACAGAGAGGAGAAACAAACAACAAGAAACAAAGATATCCAATGGCACT 3799
Db 1524 TCAGAGAAGAACACAGAGAGGAGAAACAAACAACAAGAAACAAAGATATCCAATGGCACT 1583

QY 3800 TGATCTTAACCTTCACAGCTCAGTTTAAACACCAAGTTGATGATCAAGAGGAGAGAGAAACAC 3859
Db 1584 TGATCTTAACCTTCACAGCTCAGTTTAAACACCAAGTTGATGATCAAGAGGAGAGAGAAACAC 1643

QY 3860 AGGATTTCTTTGGAATCGGATTAGATGCTTCAAAAGCTAATGAGTAGAGGAGAGAAACAGGTTT 3919
Db 1644 AGGATTTCTTTGGAATCGGATTAGATGCTTCAAAAGCTAATGAGTAGAGGAGAGAAACAGGTTT 1703

QY 3920 TAAACCATACAAAGAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAACAATCC 3979
Db 1704 TAAACCATACAAAGAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAACAATCC 1763

QY 3980 TATCATTCATGTGGAACAGAAAGATCCAAACCGATGCGGTGGGAACTCAAGCTTCCAC 4039
Db 1764 TATCATTCATGTGGAACAGAAAGATCCCAACCGATGCGGTGGGAACTCAAGCTTCCAC 1823

QY 4040 ATGA 4043
Db 1824 ATGA 1827

RESULT 5
US-09-887-576-702
; Sequence 702, Application US/09887576
; Patent No. US2002014407A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 702
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-702

Query Match 26.2%; Score 1137; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 3.6e-273;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 ATCAAGAGGAGGAAGAAGAAGAAGAGCGCTTTTTCAGGCGCATTTTCATGAATTCGAATGAA 257
Db 1 ATCAAGAGGAGGAAGAAGAAGAAGAGCGCTTTTTCAGGCGCATTTTCATGAATTCGAATGAA 60

QY 258 GGATATCAAAAGAACTTAACACAAAGGCCAGCTCTTCTTCAATCTTTCTCTTCTTCTTCT 317
Db 61 GGATATCAAAAGAACTTAACACAAAGGCCAGCTCTTCTTCAATCTTTCTCTTCTTCTTCT 120

QY 318 CTAATAATATTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 377
Db 121 CTAATAATATTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

QY 378 TCCATTTATGTCAAAAGTGTGTAAATTCCTCAAGACTATATATGAGATGTTTGTTCAT 437
Db 1 TCCATTTATGTCAAAAGTGTGTAAATTCCTCAAGACTATATATGAGATGTTTGTTCAT 437
```

Db	181	TCCATTTATGTCAAAGTGTGTGTAATTCCTCAAGACTATATATAGATGTTTGTGTTTCAT	240
Qy	438	TTTCCAAATTTCAAACCTTTGFCCECATTTAGTCTTCTACCCCTTCATGCATGGTTAGCTT	497
Db	241	TTTCCAAATTTCAAACCTTTGFCCECATTTAGTCTTCTACCCCTTCATGCATGGTTAGCTT	300
Qy	498	AGCTTAATGCTGAACCTGTTGAATACGATATACGAGCTTTATGCTAAAGAAACAAACCTTA	557
Db	301	AGCTTAATGCTGAACCTGTTGAATACGATATACGAGCTTTATGCTAAAGAAACAAACCTTA	360
Qy	558	TGGGTCTAAAAAATAAAGCCCAATATAAACTATGGCCCAATAAGTTTAGTGCCATTA	617
Db	361	TGGGTCTAAAAAATAAAGCCCAATATAAACTATGGCCCAATAAGTTTAGTGCCATTA	420
Qy	618	GAGTGTGAGATAGCGCGTGTAGTGAAACGACAGAGATGCGCGTTCAATTGTTGGTGAA	677
Db	421	GAGTGTGAGATAGCGCGTGTAGTGAAACGACAGAGATGCGCGTTCAATTGTTGGTGAA	480
Qy	678	GTAGTCGTCTAGATTCCTCGGGTCCACTGATGTTTCTAGTGTATCAGACACGTGTCGAA	737
Db	481	GTAGTCGTCTAGATTCCTCGGGTCCACTGATGTTTCTAGTGTATCAGACACGTGTCGAA	540
Qy	738	ACTGCTGGAGAGATTAACGATCTTTAAGTAGTCCCACTAGATCAAGATATTTATAACGAA	797
Db	541	ACTGCTGGAGAGATTAACGATCTTTAAGTAGTCCCACTAGATCAAGATATTTATAACGAA	600
Qy	798	TTGACCTTTTTAACCTTTTCAGGTAGTCCCGGAACTCGTGGCTAGAAACAAAGAGGTT	857
Db	601	TTGACCTTTTTAACCTTTTCAGGTAGTCCCGGAACTCGTGGCTAGAAACAAAGAGGTT	660
Qy	858	GTGAACAAAGTTGATGTTTAAGATGGCAAGAATGTAACTTGGAACAAAAGCTGAATCTC	917
Db	661	GTGAACAAAGTTGATGTTTAAGATGGCAAGAATGTAACTTGGAACAAAAGCTGAATCTC	720
Qy	918	TTCAGCCACTAGTATGTTTGACATATGCGCAGTTTCTTTTGTAGCCCTCGAAATAAATTAAT	977
Db	721	TTCAGCCACTAGTATGTTTGACATATGCGCAGTTTCTTTTGTAGCCCTCGAAATAAATTAAT	780
Qy	978	AAAAAGTTTGAGGTTAAAGATAATATATAGTCGCTGAGATTTCTCCATTTCCGTAGCTTCT	1037
Db	781	AAAAAGTTTGAGGTTAAAGATAATATATAGTCGCTGAGATTTCTCCATTTCCGTAGCTTCT	840
Qy	1038	GGTCTCTTTTCTTTGTTTCATTGATCAAAAGCAATCACITTCCTTCTCTCTCTCTCTCG	1097
Db	841	GGTCTCTTTTCTTTGTTTCAITGATCAAAAGCAATCACITTCCTTCTCTCTCTCTCTCG	900
Qy	1098	ATTTCTTACTGTTTTCTTTATCCACGAAATCTGAAATTTAAAAATGGAATCTTTATCGAAT	1157
Db	901	ATTTCTTACTGTTTTCTTTATCCACGAAATCTGAAATTTAAAAATGGAATCTTTATCGAAT	960
Qy	1158	CCAGCTGATTTTGTTCCTTTTCATTGATCATCTCTCTAAAGGTACTTTAAGATTGATTTA	1217
Db	961	CCAGCTGATTTTGTTCCTTTTCATTGATCATCTCTCTAAAGGTACTTTAAGATTGATTTA	1020
Qy	1218	TTGTCACTGGTCTTCTTATCTTTTGATGAATAAATCTGACTTGATGTTTTTTGTTTTGTG	1277
Db	1021	TTGTCACTGGTCTTCTTATCTTTTGATGAATAAATCTGACTTGATGTTTTTTGTTTTGTG	1080
Qy	1278	GATTAGTGGAAATTTTGTAAAGAGAAGATCTGAAGTTGTGTAGAGAGCTTAGTCAATG	1334
Db	1081	GATTAGTGGAAATTTTGTAAAGAGAAGATCTGAAGTTGTGTAGAGAGCTTAGTCAATG	1137

## RESULT 6

```

US09-938-842A-4192
; Sequence 4192, Application US/09938842A
; Patent No. US2002160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krebs, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
US09-938-842A-4192

```

	Query Match	7.2%	Score 312.6;	DB 9;	Length 865;
	Best Local Similarity	95.8%;	Pred. No. 3e-67;		
	Matches 321;	Conservative	0;	Mismatches 14;	Indels 0;
				Gaps	0;
Qy	114	AGGAATTTGTTTGTATCTTTTCAGGAGGCTTTGTTTCAGAGGTCTCAAGAGCATATC	173		
Db	335	AAAGACCTGTCTCTCAAAACGAGACGAGGAGGCTTTGTTTGAGAGGTCTCAAGAGCATATC	276		
Qy	174	AAAGACATATTTAGGGAGCAGCTCGAATCAAGGAGGAGGAAGAAGAGGAGGCTTTTTC	233		

	Query Match	6.7%	Score 289;	DB 11;	Length 479;
	Best Local Similarity	73.8%;	Pred. No. 1.6e-61;		
	Matches 479; Conservative	0;	Mismatches	0;	Indels 170; Gaps 2
Qy	2241	ATAGCTATTCTCTCCACGCCCTAAGCGTAAACCAACAATCCTTTATCTCTCGAAAGACG	2300		
Db	1	ATAGCTATTCTCTCTCCACGCCCTAAGCGTAAACCAACAATCCTTTATCTCTCGAAAGACG	60		
Qy	2301	GGAAGTGGAACGATCCTTTATGTCAAACACGGGTGTGAATGATGAAAAAGATCCCTTGGA	2366		
Db	61	GGAAGTGGAACGATCCTTTATGTCAAACACGGGTGTGAATGATGAAAAAGATCCCTTGGA	120		
Qy	2361	TCAGAAAAAGTGTCCGATCCTTGAGGTGATTTTTTCATGGTGCATATGGCATCTTTTTTGCAGTG	2420		



Db 121 TCAGAAAAAGTGTGCGATCCTG----- 142  
QY 2421 TGTACATTGCTCTCATGTATTATACAGATTGTGCTGCTGTTTATAGATGGCCAA 2480  
Db 143 -----AGATGGCCAA 153  
QY 2481 GAAGATCGACACAATCAAAAGCTGAAGAGAAACTCTCCAGAGACAACTGTTTCAGAT 2540  
Db 154 GAAGATCGACACAATCAAAAGCTGAAGAGAAACTCTCCAGAGACAACTGTTTCAGAT 213  
QY 2541 TGTTCCTACATCATGATCTCTGCTGCTGATCCTCATCAATAAAGTTGTATAGAGACA 2600  
Db 214 TGTTCCTACATCATGATCTCTGCTGCTGATCCTCATCAATAAAGTTGTATAGAGACA 273  
QY 2601 TCAAGCGAAGCACTTTCGCGAGTTCTTTCGCTTTCACGGGAAGAGGTAATAAAACAATCTT 2660  
Db 274 TCAAGCGAAGCACTTTCGCGAGTTCTTTCGCTTTCACGGGAAG----- 316  
QY 2661 TCATTGCTATTGAGGTTTAAAGCAATAGTACTTTTCATGAACATAAACCGTGGGG 2720  
Db 317 ----- 316  
QY 2721 AATAACAGGGAAGTGCAGAAATACAGGGTAAGAAAGGTCAAACTCAGATTTGAATGCAA 2780  
Db 317 -----AGGAGTCAAGATACAGGGTAAGAAAGGTCAAACTCAGATTTGAATGCAA 370  
QY 2781 AATCTCTGAAAAACGGTAATAGCAAGGACCTCAGACTTATCCGATGCATATCCCTGTGC 2840  
Db 371 AATCTCTGAAAAACGGTAATAGCAAGGACCTCAGACTTATCCGATGCATATCCCTGTGC 430  
QY 2841 TAGTCCATTGGGAGCTCAATAAAGTTCTCTATCATCTCTCTC 2889  
Db 431 TAGTCCATTGGGAGCTCAATAAAGTTCTCTATCATCTCTCTC 479

RESULT 9

US-09-934-455-119  
; Sequence 119, Application US/09934455  
; Publication No. US20030121070A1  
; GENERAL INFORMATION:  
; APPLICANT: Creelman, Robert  
; APPLICANT: Dubell, Arnold  
; APPLICANT: Heard, Jacqueline  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Keddie, James  
; APPLICANT: Pilgrim, Marsha  
; APPLICANT: Reuber, Lynne  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Pineda, Omaira  
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
; FILE REFERENCE: MBI-0025  
; CURRENT APPLICATION NUMBER: US/09/934,455  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/227439  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: MBI-0022  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: MBI-0023  
; PRIOR FILING DATE: 2001-04-17  
; NUMBER OF SEQ ID NOS: 516  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 119  
; LENGTH: 2526  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (338)..(2275)  
; OTHER INFORMATION: G680  
US-09-934-455-119

Query Match 1.9%; Score 83.2; DB 11; Length 2526;  
Best Local Similarity 50.9%; Pred. No. 1.3e-09;  
Matches 233; Conservative 0; Mismatches 213; Indels 12; Gaps 1;  
QY 2847 CATGGGAGCTCAATAACAAGTTCTCTATCATCATCTCTCTCAGAGCCAGATGATCATC 2906  
Db 1137 CATCTGCTACTACTACAGCTTCTCATCAAGCGTTTCCAGCTTGTCTTACAGAGATGATT 1196  
QY 2907 CCCACACAGTTCAGAGAGATTATAGTCTGTTTCTTAATCATATAATGTCAACCTTTTAC 2966  
Db 1197 ACCGTTCGTTTCTCCAGATATCATCTACTTTCTCCAATCTTATATGTCAACTCTCCTAC 1256  
QY 2967 AAACACCGGCTCTTTATATCTGCGGAACTTTCGCTCATCATTTTGGCGCTCCCG----- 3020  
Db 1257 AGAATCTGCGAGCTCATGCTGAGTACATTGCTGCTTGGTCTGGCTTATGCGAGTG 1316  
QY 3021 -----ATTCTAGTGGTGGCTCACCTGTTTCCAGGAACTCACCTCCGAATCTGGCTGCCA 3074  
Db 1317 TCGGAAATTCGGTGATTTCATCAACCCCAATGAGCTCTTCTCTCCAAGTATAAATGCCA 1376  
QY 3075 TGGCGCGAGCCACTGTTGCGAGCTGCTAGTGTCTTGGTGGCTGCCAATGATTTACCTT 3134  
Db 1377 TTGCGCTGCTACAGTAGCTGCTGCAACTGCTTGGTGGCTTCTCATGGACTTCTTCTCTG 1436  
QY 3135 TATGTGCTCTCTTAGTTAGTTCAGGTGTTTCACTAGTATCATCTCCATCTCTTTGGACCAT 3194  
Db 1437 TATGGCTCCAGCTCAATACATGTGTTCCATCTTCAACTGTTGCGATTCCAACTCCAG 1496  
QY 3195 CATGTGATGTAGAGTACACAAAGCAAGCACTTTTCAACATGTTCTGTGCGAGCCGAG 3254  
Db 1497 CAATGACTGAATGGATACCGTTTGAATACTCAACCGTTTGAGAAACAACACAGCTC 1556  
QY 3255 ACAGAAACACTCCGAGGATCAAAAGGCTCGATCTTCA 3292  
Db 1557 TGCAAGATCAAACTTGGCTTCGAAATCTCCAGCTTCA 1594

RESULT 10

US-10-278-536-199  
; Sequence 199, Application US/10278536  
; Publication No. US20030131386A1  
; GENERAL INFORMATION:  
; APPLICANT: Samaha, Raymond  
; APPLICANT: Heard, Jacqueline  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Reuber, Lynne  
; APPLICANT: Riechmann, Jose-Luis  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Keddie, James  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Pilgrim, Marsha  
; APPLICANT: Adam, Luc  
; APPLICANT: Broun, Pierre  
; TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES  
; FILE REFERENCE: MBI-011  
; CURRENT APPLICATION NUMBER: US/10/278,536  
; CURRENT FILING DATE: 2002-10-22  
; PRIOR APPLICATION NUMBER: 60/125,814  
; PRIOR FILING DATE: 1999-03-23  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 199  
; LENGTH: 2526  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: G680  
US-10-278-536-199

Query Match 1.9%; Score 83.2; DB 15; Length 2526;  
Best Local Similarity 50.9%; Pred. No. 1.3e-09;



	Matches	233;	Conservative	0;	Mismatches	213;	Indels	12;	Gaps	1;
Qy	2847	CATTGGGGAGCTCAATAACAAAGTTCCTCTATCACATCTCTCTTCAGAGCCAGATAGTCATC	2906							
Db	1137	CATCTGCTACTACTACAGCTTCTCATCAAGCGTTTCCAGCTTGTCATTACAGGATGATT	1196							
Qy	2907	CCACACAGTTGTCAGGAGATTATCAGTCGTTTCTCTAATCATATATGTCAACCCCTTTTAC	2966							
Db	1197	ACCGTTTGGTTTCTCCAGATATCATCTACTTTTCTCCAAATCTTATTATGTCAACTCTCTCTAC	1256							
Qy	2967	AAACACCGGCTCTTTTATCTGCGCCAACTTTTCGGCTCATCATTTTGGCTCTCCG-----	3020							
Db	1257	AGAATCTGCAGCTCATGTCGAGCTACATTCCTCGTCTCGGTCTGCGCTTATGCGAGTG	1316							
Qy	3021	-----ATTCTTAGTGGTGGCTCACTGTTCCAGGGAATCACTCCGAATCTGGCTGCCA	3074							
Db	1317	TCCGGAAATCTGGTGAATTCATCAACCCCAATGAGCTCTTCTCTCCAAGTATAACTGCCA	1376							
Qy	3075	TGGCCGAGCCACATGTTTGCAGCTGCTAGTGTCTGGTGGGTGCCAAATGGATATTATACCTT	3134							
Db	1377	TTGCGCTGCTACAGTAGCTGCTGCAACTGCTTGGTGGGCTTCTCATGAGACTTCTTCCTG	1436							
Qy	3135	TATGTGCTCTCTTAGTTTCAGGTGGTTTCACTAGTCAATCTCTCAATCTACTTTTGGACCAT	3194							
Db	1437	TATGCGTCCAGCTCCCAATAACATGTGTTCATTTCTCAACTGTTTGCAGTTCCTCAACTCCAG	1496							
Qy	3195	CATGTGATGTAGAGTACACAAAAACCAAGCACTTTTACAACATGGTTCTGTGCAGAGCCGAG	3254							
Db	1497	CAATGACTGAATGATATCCGTTGAAATACTCAACCGTTTGAAACAAAAACACAGCTC	1556							
Qy	3255	AGCAAGAACTCTCGAGGGCATCAAAGGCTCGATCTTCA	3292							
Db	1557	TGCAAGATCAAACTTGGCTTCGAAATCTCGAGCTTCA	1594							

RESULT 11  
 US-09-938-842A-1116  
 ; Sequence 1116, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; FILE REFERENCE: SAME, AND METHODS OF USE  
 ; FILE REFERENCE: S01P1300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 1116  
 ; LENGTH: 1938  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-1116

	Query Match	1.9%	Score 81.6;	DB 10;	Length 1938;
	Best Local Similarity	50.7%;	Pred. No. 2.8e-09;		
	Matches 233;	Conservative	0;	Mismatches 214;	Indels 12; Gaps 1;
QY	2847	CATTGGGACCTCAATTAACAAAGTTCCTATCATCCTCTTCAGAGCCAGATAGTCAATC	2906		
Db	800	CATCTGCTACTACTACAGCTTCTCATCAAGCGTTTCCAGCTTGTCATTCACAGGATGATT	859		
QY	2907	CCACACAGTTGCAGGAGATTATCAGTCGTTTCTTATATCATATAATGTCAACCCCTTTTAC	2966		
Db	860	ACGGTTGTTTCCAGATATCATCTACTTTCTCCAATCTTATTATGTCAACTCTCCTAC	919		

```

RESULT 12
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeuten
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated gel
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

```

Query Match	1.5%;	Score 66.4;	DB 12;	Length 3673778;
Best Local Similarity	47.2%;	Pred. No. 0.0031;		
Matches 331;	Conservative 0;	Mismatches 366;	Indels 5;	Gaps 4;
Qy	947	GTITCTTTTTGTAGCCTCGAAATAAATAAATAAAGTTTGAAGTTAAAGATATAATTATAG	1006	
Db	56481	GATTTTGTTTTTTTTTTAGGTTTATATATGGTAAGAAATTTTAGTTTTTCGTATTTTTTAA	56540	
Qy	1007	TGGCTCAGAGATTTCTCCAAITTCGTAGCTTCGTCTCTTTCTTTCTT-TGTTTCAATGCATCAA	1065	
Db	56541	TTTTATATTTTTTATTTTTTTGATTAATTTTTTATTTTTTTTTTGAGGTAGTTATAGG	56600	
Qy	1066	AAGCAAAATCACCTCTCTTCTTCTTCTCTCGAATTTCTTACTGTGTTTTCTTATCCAAACGAA	1125	
Db	56601	TTTTGAAGATAATTGATGTTATATAATTTGATTATATGTTATATATGTAATATTAGTGAATAT	56660	
Qy	1126	ATCTGGAATTAATAATGGAACTTTATCGAATCCAAAGCTGAATTTTGTTCTTTTCATTTGAA	1185	
Db	56661	TTATTTGATGCTGTGTTGTTTTTATAGGTATGAGGTTTATTTTTTGATGATATATTTATTTAA	56720	
Qy	1186	TCATCTCTCTAAAGGTACTTAAAGATTGAATTTATTTGTCATGCTCTTCTTATTTGTTTCATG	1245	
Db	56721	TAAATTTTTTATTTTTTTTGAATAATTTTAAAT-TAGTGAATGTGTTATTTATTTTTTTTT	56779	
Qy	1246	AATAACTTGACTTGATGTTTTTTTTGTTGTGGATTAGTGGAAATTTTGTAAAGAGAAGAT	1305	



; NUMBER OF SEQ ID NOS: 5379

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2003, 01:20:08 ; Search time 8553.69 Seconds  
(without alignments)  
12343.052 Million cell updates/sec

Title: US-10-084-553-1

Perfect score: 4344

Sequence: 1 Gcagtggttcacttaacaaga.....actcataagtaagcacaataa 4344

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	735.2	16.9	740	28	AQ958316 LERAW81TF
C 2	430	9.9	430	9	AV809889
C 3	428	9.9	439	9	AV792310
C 4	425	9.8	426	9	AV804753

5	410.6	9.5	436	28	AQ958316
6	397.4	9.1	421	29	CC179314
7	397	9.1	421	29	CC179316
8	395.8	9.1	515	9	AA394931
C 9	393	9.0	393	9	AV814102
C 10	384	8.8	396	9	AV801368
C 11	366.4	8.4	393	28	AQ958317
C 12	360	8.3	360	9	AV784525
C 13	356.2	8.2	423	9	AV807633
C 14	339.4	7.8	580	9	AV827752
C 15	327.6	7.5	415	9	AV787515
C 16	320.6	7.4	415	9	AV811181
C 17	318.4	7.3	628	9	AV823536
C 18	315.6	7.3	403	9	AV799191
C 19	314.6	7.2	569	9	AV556228
C 20	306.6	7.1	396	9	AV545439
C 21	306.6	7.1	410	9	AV785900
C 22	306.6	7.1	415	9	AV803584
C 23	306.6	7.1	415	9	AV818664
C 24	306.6	7.1	418	9	AV799094
C 25	306.6	7.1	540	9	AV566656
C 26	306.6	7.1	685	9	AV784812
C 27	303.4	7.0	418	9	AV811618
C 28	302.6	7.0	391	9	AV797930
C 29	302.6	7.0	420	9	AV810366
C 30	302.6	7.0	577	9	AV534474
C 31	302	7.0	330	14	T03973
C 32	301.6	6.9	434	9	AV802892
C 33	299.6	6.9	422	9	AV816966
C 34	299.6	6.9	426	9	AV791925
C 35	298	6.9	422	9	AV819444
C 36	298	6.9	427	9	AV791998
C 37	298	6.9	436	9	AV819310
C 38	296.4	6.8	572	9	AV567367
C 39	296	6.8	296	9	AV565983
C 40	295.4	6.8	487	9	AV92931
C 41	283.6	6.5	438	9	AV520422
C 42	279.6	6.4	653	9	AV520014
C 43	278.6	6.4	605	9	AV520854
C 44	277.6	6.4	452	9	AV794212
C 45	269.2	6.2	403	9	AV787122

#### ALIGNMENTS

RESULT 1  
AQ958316/c

LOCUS LERAW81TF LERA Arabidopsis thaliana genomic clone LERAW81, genomic survey sequence.  
DEFINITION

ACCESSION AQ958316  
VERSION AQ958316.1 GI:6786017

KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 740)  
AUTHORS Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T., Feldblum, T., Liang, F., Creasy, T. and Fraser, C.M.

TITLE

Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms  
JOURNAL Unpublished  
COMMENT Contact: Xiaoying Lin  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: at@tigr.org

For additional information, see <http://www.tigr.org/tdb/at/at.html>

Seq primer: TF  
 Class: shotgun.  
 Location/Qualifiers  
 1..740

# FEATURES

source  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Landsberg erecta"  
 /db\_xref="taxon:3702"  
 /clone\_lib="LERAW81"  
 /note="Organ: Leaf; Vector: pHOS1, Total genomic DNA was sheared to 0.9-1 Kbp before ligation."  
 BASE COUNT 184 a 176 c 178 g 202 t  
 ORIGIN

Query Match 16.9%; Score 735.2; DB 28; Length 740;  
 Best Local Similarity 99.6%; Pred. No. 3.2e-107;  
 Matches 737; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2887 TTCAGACCCAGATAGTCATCCACACAGATTGCAGGAGATTATCAGTCGTTTCTTAATCA 2946  
 Db 740 TTCAGACCCAGATAGTCATCCACACAGATTGCAGGAGATTATCAGTCGTTTCTTAATCA 681

QY 2947 TATAATGTCACACCCCTTTTACAAACACCGGCTCTTTTATCTGCGCAACTTTCGCTCATC 3006  
 Db 680 TATAATGTCACACCCCTTTTACAAACACCGGCTCTTTTATCTGCGCAACTTTCGCTCATC 621

QY 3007 ATTTGGCCCTCCCGATTTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACTCCGAATCT 3066  
 Db 620 ATTTGGCCCTCCCGATTTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACTCCGAATCT 561

QY 3067 GGCTGCCATGGCGCAGCCACTGTTGAGCTGCTAGTGTGTTGGTGGCTGCCAATGGATT 3126  
 Db 560 GGCTGCCATGGCGCAGCCACTGTTGAGCTGCTAGTGTGTTGGTGGCTGCCAATGGATT 501

QY 3127 ATTACCTTTATGTGCTCTCTTAGTTCAGGTGGTTTCACTAGTCATCTCCATCTACTTT 3186  
 Db 500 GTTACCTTTATGTGCTCTCTTAGTTCAGGTGGTTTCACTAGTCATCTCCATCTACTTT 441

QY 3187 TGGACCATCATGTGATGTAGAGTACACAAAGCAAGCACTTTCAACATGGTTCTGTGCA 3246  
 Db 440 TGGACCATCATGTGATGTAGAGTACACAAAGCAAGCACTTTCAACATGGTTCTGTGCA 381

QY 3247 GAGCCGAGACGACAACTCCGAGGATCAAAAGCTCGATCTTCACTGGACTCAGAGGA 3306  
 Db 380 GAGCCGAGACGACAACTCCGAGGATCAAAAGCTCGATCTTCACTGGACTCAGAGGA 321

QY 3307 TGTGAAATTAAGAGTAAACCAAGTTTGTCTAGTCAGCGCTTCTGCAACACCTCGAGGTGA 3366  
 Db 320 TGTGAAATTAAGAGTAAACCAAGTTTGTCTAGTCAGCGCTTCTGCAACACCTCGAGGTGA 261

QY 3367 TGCAAGGGTTTCAGATGAGGACGAGACAGAAACAAAGTTGACCGGTCTCTCGTGTGGCTC 3426  
 Db 260 TGCAAGGGTTTCAGATGAGGACGAGACAGAAACAAAGTTGACCGGTCTCTCGTGTGGCTC 201

QY 3427 AATACCTCCGTCAGTGTATGATGTTGAGGCGGATGCATCAGAAAGGCAAGAGATGG 3486  
 Db 200 AATACCTCCGTCAGTGTATGATGTTGAGGCGGATGCATCAGAAAGGCAAGAGATGG 141

QY 3487 CACCAATGGTGGAGTGAAGAAACCAATGAAGACACTAATAAACCTCAAACTTCAGAGTC 3546  
 Db 140 CTCCAATGGTGGAGTGAAGAAACCAATGAAGACACTAATAAACCTCAAACTTCAGAGTC 81

QY 3547 CAATGACGCGCGAGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGTCTGA 3606  
 Db 80 CAATGACGCGCGAGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGTCTGA 21

QY 3607 CGAGGTACTTACTTGGACTA 3626  
 Db 20 CGAGGTACTTACTTGGAGTA 1

RESULT 2

AV809889/c

LOCUS

DEFINITION

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

AV809889

```

source
1. .426
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-40-G07"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
136 a 75 c 82 g 133 t
BASE COUNT

```



This is single pass sequence recovered from the left border of TDNA.

FEATURES	Class: TDNA tagged.	Location/Qualifiers	
source	1. .421	/organism="Arabidopsis thaliana"	
		/mol_type="genomic DNA"	
		/strain="Columbia 0"	
		/db_xref="taxon:3702"	
		/clone="SALK_067774.47.95.x"	
		/clone_lib="Arabidopsis thaliana TDNA insertion lines"	
		/notes="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at <a href="http://signal.salk.edu/tdna_protocols.html">http://signal.salk.edu/tdna_protocols.html</a> "	
BASE COUNT	136 a	44 c	91 g 150 t
ORIGIN			
	Query Match	9.1%	Score 397.4; DB 29; Length 421;
	Best Local Similarity	98.1%;	Pred. No. 2.3e-53;
	Matches 413;	Conservative 0;	Mismatches 6; Indels 2; Gaps 1;
QY	1102	CTTACTGTTTCTTATCCAAAGAAATCTGGAATTTAAAAATGGAATCTTTATCGAATCCAA	1161
DB	1	CTTACTGTTTCTTATCCAAAGAAATCTGGAATTTAAAAATGGAATCTTTATCGAATCCAA	60
QY	1162	GCTGATTTTCTTTCTTTTCATTGAATCATCTCTCTAAAGGTACTTTAAGATTGATTATTGT	1221
DB	61	GCTGATTTTCTTTCTTTTCATTGAATCATCTCTCTAAAGGTACTTTAAGATTGATTATTGT	120
QY	1222	CATGGTCTTTCTTATGCTTTTGATGAATAACTTGACTTGATTGTTTTTGTGTTGTTGATT	1281
DB	121	CATGGTCTTTCTTATGCTTTTGATGAATAACTTGACTTGATTGTTTTTGTGTTGATT	180
QY	1282	AGTGGAAATTTGTAAAGAGAAGATCTCAAGTTGTGTAGAGAGCTTAGTGTATGGAGACAA	1341
DB	181	AGTGGAAATTTGTAAAGAGAAGATCTCAAGTTGTGTAGAGAGCTTAGTGTATGGAGACAA	240
QY	1342	ATTTCGTCGTGAGAAGATCTCGTTATTTAAGGTAAATTA -- CTAAATTTTAGGGGGAAGAT	1399
DB	241	ATTTCGTCGTGAGAAGATCTCGTTATTTAAGGTAAATTA -- CTAAATTTTAGGGGGAAGAT	300
QY	1400	GATTGTTTTTAGTGTCAAAGATTTCGAAATTTTAATGAACTTGATATAGACTCGGAAGCC	1459
DB	301	GATTGTTTTTAGTGTCAAAGATTTCGAAATTTTAATGAACTTGATATAGACTCGGAAGCC	360
QY	1460	ATATACGATAACAAAGCAACGCTGAAAGGTCGACTGAGCAAGAACATAATAGATTCATTGA	1519
DB	361	ATATACGATAACAAAGCAACGCTGAAAGGTCGACTGAGCAAGAACATAATAGATTCATTGA	420
QY	1520	A 1520	
DB	421	A 421	

RESULT 7  
CC179316  
LOCUS  
DEFINITION  
CC179316 421 bp DNA linear GSS 02-MAY-2003  
SALK\_067780.51.35.x Arabidopsis thaliana TDNA insertion lines  
Arabidopsis thaliana genomic clone SALK\_067780.51.35.x, genomic  
survey sequence.

sequence.

CC179316  
CC179316.1 GI:30317867  
GSS.  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 421)  
REFERENCE  
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab

C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J. and Ecker, J.R.  
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome  
Unpublished  
Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGNAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of TDNA.

```

FEATURES
source
Class: TDNA tagged
Location/Qualifiers
1. 421
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GALK 067780.51.35.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna\_protocols.html"
135 a 44 c 92 g 150 t
BASE COUNT
ORIGIN

```

Query Match	9.1%;	Score 397;	DB 29;	Length 421;
Best Local Similarity	99.5%;	Pred. No. 2.6e-53;		
Matches 419;	Conservative 0;	Mismatches 0;	Indels 2;	Gaps 2;
QY	1102	CTTACTGTTTCTCTTATCCACGAAATCTCGAATTTAAAATGGAATCTTTATCGAAATCCAA	1161	
DB	1	CTTACTGTTTCTCTTATCCACGAAATCTCGAATTTAAAATGGAATCTTTATCGAAATCCAA	60	
QY	1162	GCTGATTTTGTCTTCTTTCATTGAATCATCTCTCTAAAGGTACTTAAGATTGATTTATT - G	1220	
DB	61	GCTGATTTTGTCTTCTTTCATTGAATCATCTCTCTAAAGGTACTTAAGATTGATTTATTGG	120	
QY	1221	TCATGCTCTTCTTATTGTTTGATGAATAACTTGACTTGCATTTGTTTTGTTTGTGCGAT	1280	
DB	121	TCATGCTCTTCTTATTGTTTGATGAATAACTTGACTTGCATTTGTTTTGTTTGTGCGAT	180	
QY	1281	TAGTGAATTTTGTAAAGAGAAGATCTCGAAGTTGTGTAGAGAGCTTTAGTGTGGAGACA	1340	
DB	181	TAGTGAATTTTGTAAAGAGAAGATCTCGAAGTTGTGTAGAGAGCTTTAGTGTGGAGACA	240	
QY	1341	AATTCGCTCTGGAGAAGATCTGGTTATTAAAGTTAAATT - AACTTAAATTTTAGGGGGAAGAT	1399	
DB	241	AATTCGCTCTGGAGAAGATCTGGTTATTAAAGTTAAATTAAACTTAAATTTTAGGGGGAAGAT	300	
QY	1400	GATTGTTTTAGGTGTCAAAAGATTGAGAAATTTTAAATGAAACTTGATATAGACTCGGAAGCC	1459	
DB	301	GATTGTTTTAGGTGTCAAAAGATTGAGAAATTTTAAATGAAACTTGATATAGACTCGGAAGCC	360	
QY	1460	ATATACGATAACAAAGCAACGTGAAAGGTGGACTGAGGAAGAACATAAATAGATTCATTGA	1519	
DB	361	ATATACGATAACAAAGCAACGTGAAAGGTGGACTGAGGAAGAACATAAATAGATTCATTGA	420	
QY	1520	A 1520		
DB	421	A 421		

RESULT 8					
AA394931	AA394931	515 bp	mRNA	linear	EST 30-OCT-1997
LOCUS	26812	Lambda-PRL2	Arabidopsis	thaliana	CDNA clone 109K9XP 3', mRNA
DEFINITION					



sequence.	347	TATTTTCATCTGATCTGTTGTTGTACTCTGTTTAAAGTTTTCAGAGCCACTGCTACAT	406
ACCESSION	AA394931		
VERSION	AA394931.1	GI:2048226	
KEYWORDS	EST.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh L., Ohlrogge,J., Raikheil,N., Somerville,S., Thomashow,M., Retzel E. and Somerville,C.		
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones		
JOURNAL	Plant Physiol.	106, 1241-1255 (1994)	
MEDLINE	95148729		
PUBMED	7846151		
COMMENT	Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan State University MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI Tel: 517-353-0854 Fax: 517-353-9168 Email: 22313tcn@ibm.cl.msu.edu Seq primer: M13 -21 dye primer.		
FEATURES	Location/Qualifiers		
source	1. 515		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/strain="var columba"		
	/db_xref="taxon:3702"		
	/clone_lib="109K9XP"		
	/clone_lib="Lambda-PRL2"		
	/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA. 130 a 118 c 96 g 152 t 19 others		
BASE COUNT	130 a 118 c 96 g 152 t 19 others		
ORIGIN			
Query Match	9.1%;	Score 395.8;	DB 9; Length 515;
Best Local Similarity	91.9%;	Pred. No. 3.9e-53;	
Matches 431;	Conservative 0;	Mismatches 36;	Indels 2; Gaps 2;
QY	3749	ACACAGAGAGGAAGAAACAACAACAAGAACAAAGATATCCATGGCACTTCTTAA	3808
Db	49	ACCCCGGGGGAGGTCCACCNCACNCAGCAAGNTTTCCTCAANGCCTTG-TCTTAA	107
QY	3809	CTTCACAGCTCAGTTAAACACCGATTGATGATCAAGAGGAGAGAGAAACACAGGATTTCT	3868
Db	108	CTTCNCGGCCAGTTAAACCCCGTTGTGACCCAGGGGNGAGAGNCCAGGA-TTCT	166
QY	3869	TGGAATTCGGATTAGATGCTTCAAAGCTTAATGAGTAGAGGAAGAACAGGTTTTAAACCAT	3928
Db	167	TGGAACGGATTAGATGCTCAAAGCTTAATGAGTAGAGGAAGAACAGGTTTTAAACCAT	226
QY	3929	CAAAAGATGTTCCATGGAGCCCAAGAAAGTAGAATCCTCAACAACAATCCTATCATTTCA	3988
Db	227	CAAAAGATGTTCCATGGAGCCCAAGAAAGTAGAATCCTCAACAACAATCCTATCATTTCA	286
QY	3989	TGTGGAACAGAAAGATCCCAACCGATGCGGTTGGAAACTCAAGCTTCCACATGAGACTC	4048
Db	287	TGTGGAACAGAAAGATCCCAACCGATGCGGTTGGAAACTCAAGCTTCCACATGAGACTC	346
QY	4049	TATTTTCATCTGATCTGTTGTTGTACTCTGTTTAAAGTTTTCAGAGCCACTGCTACAT	4108

QY 3935 ATGTTCCATGAAGCCAAAGAAAGTAGAATCCTCAACAAACATCTATCATTTGATGGA 3994  
 |||||  
 Db 273 ATGTTCCATGAAGCCAAAGAAAGTAGAATCCTCAACAAACATCTATCATTTGATGGA 214  
 |||||  
 QY 3995 ACAGAAAGATCCCAACCGGATGGGTTGGAACCTCAAGCTTCCACATGAGACTCTATTTT 4054  
 |||||  
 Db 213 ACAGAAAGATCCCAACCGGATGGGTTGGAACCTCAAGCTTCCACATGAGACTCTATTTT 154  
 |||||  
 QY 4055 CATCTGATCTGTTGTTGTAATCTGTTTAAAGTTTCAAGACCACTGCTACATTTTCTT 4114  
 |||||  
 Db 153 CATCTGATCTGTTGTTGTAATCTGTTTAAAGTTTCAAGACCACTGCTACATTTTCTT 94  
 |||||  
 QY 4115 TTCTTTTGAAGCCTTGTATTTGTTTCCCTGTCATGCTTCTGTAACATTTGACTC 4174  
 |||||  
 Db 93 TTCTTTTGAAGCCTTGTATTTGTTTCCCTGTCATGCTTCTGTAACATTTGACTC 34  
 |||||  
 QY 4175 TGTATTATTCAACAAATCATAAACTGTTTAATC 4207  
 |||||  
 Db 33 TGTATTATTCAACAAATCATAAACTGTTTAATC 1

RESULT 10  
 AV801368/c 396 bp mRNA linear EST 29-MAR-2002  
 LOCUS AV801368 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-27-N02 3',  
 mRNA sequence.

ACCESSION AV801368.1 GI:19835353

VERSION AV801368

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 396)  
 AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,  
 Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,  
 Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.  
 and Shinozaki,K.  
 TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 JOURNAL Unpublished  
 COMMENT Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msekiorc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
 and XhoI was ligated to modified Lambda PUC-1 vector (Carninci et  
 al., submitted for publication) digested with BamHI and SalI. This  
 clone is in a modified pluscript vector. Please visit our web  
 site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
 details.

FEATURES Location/Qualifiers  
 1..396  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL09-27-N02"  
 /dev\_stage="plants at various developmental stages from  
 germination to mature seeds"  
 /lab\_host="DH10B"  
 /clone\_lib="RAFL9"  
 /note="Site 1: BamHI; Site 2: SalI; subjected to  
 dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
 hr) treatments"

BASE COUNT 133 a 68 c 74 g 121 t

Query Match 8.8%; Score 384; DB 9; Length 396;

Best Local Similarity 99.7%; Pred. No. 3.1e-51;  
 Matches 395; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 3823 TAACACAGTTGATGATCAAGAGGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTAG 3882  
 |||||  
 Db 396 TAACACAGTTGATGATCAAGAGGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTAG 337  
 |||||

QY 3883 ATGCTTCAAAAGCTAAATGAGTAGAGAGAGAACAGGTTTTTAAACCATACAAAAGATGTTCCA 3942  
 |||||  
 Db 336 ATGCTTCAAAAGCTAAATGAGTAGAGAGAGAACAGGTTTTTAAACCATACAAAAGATGTTCCA 277  
 |||||

QY 3943 TGAAGCCAAAGAAAGTAGAATCCTCAACAACTCTATCATTTGATGGAACAGAAAG 4002  
 |||||  
 Db 276 TGAAGCCAAAGAAAGTAGAATCCTCAACAACTCTATCATTTGATGGAACAGAAAG 217  
 |||||

QY 4003 ATCCCAACCGATGCGGTTGGAACCTCAAGCTTCCACATGAGACTCTATTTTCATCTGAT 4062  
 |||||  
 Db 216 ATCCCAACCGATGCGGTTGGAACCTCAAGCTTCCACATGAGACTCTATTTTCATCTGAT 157  
 |||||

QY 4063 CTGTTGTTGTAATCTGTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTCTTTT 4122  
 |||||  
 Db 156 CTGTTGTTGTAATCTGTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTCTTTT 98  
 |||||

QY 4123 GAGGCCCTTTGATTTGTTTCTTGTCCATAGTCTTCTGTAACATTTGACTCTGATTAT 4182  
 |||||  
 Db 97 GAGGCCCTTTGATTTGTTTCTTGTCCATAGTCTTCTGTAACATTTGACTCTGATTAT 38  
 |||||

QY 4183 TCAACAAATCATAACTGTTTAAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4218  
 |||||  
 Db 37 TCAACAAATCATAACTGTTTAAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2

RESULT 11  
 AQ958317

LOCUS AQ958317

DEFINITION AQ958317 393 bp DNA linear GSS 28-JAN-2000

ACCESION AQ958317

VERSION AQ958317.1 GI:6786018

KEYWORDS GSS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE Arabidopsis thaliana; Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Streptophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 393)  
 AUTHORS Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T.,  
 Feldblyum,T., Liang,F., Creasy,T. and Fraser,C.M.  
 TITLE Genomic survey sequencing of Landsberg erecta ecotype of  
 Arabidopsis thaliana and identification of sequence-based  
 polymorphisms  
 JOURNAL Unpublished  
 COMMENT Contact: Xiaoying Lin  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: at@tigr.org  
 For additional information, see <http://www.tigr.org/tdb/at/at.html>  
 Seq primer: TR  
 Class: shotgun.

FEATURES Location/Qualifiers  
 1..393  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Landsberg erecta"  
 /db\_xref="taxon:3702"  
 /clone="LERAW81"  
 /clone\_lib="LEA"  
 /note="Organ: Leaf; Vector: PHOS1; Total genomic DNA was  
 sheared to 0.9-1 Kbp before ligation."

BASE COUNT 90 a 108 c 84 g 111 t

Query Match 8.4%; Score 366.4; DB 28; Length 393;  
 Best Local Similarity 98.4%; Pred. No. 2e-48;  
 Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2766 CAGATTGTAATGCAAAATCTCTGGAACCGTAAATGAGCAAGGACCTCAGACTTATCCGA 2825  
 DB 13 CAGATTGTAATGCAAAATCTCTGGAACCGTAAATGAGCAAGGACCTCAGACTTATCCGA 72  
 QY 2826 TGCATATCCCTGTGCTAGTGGGAGCTCAATAAAGTTCCTATCACATCCTC 2885  
 DB 73 TGCATATCCCTGTGCTAGTGGGAGCTCAATAAAGTTCCTATCACATCCTC 132  
 QY 2886 CTTGAGAGCAGATAGTATATCCCAACAGTTCGAGGAGTATATCAGTTCCTTAATC 2945  
 DB 133 CTTGAGAGCAGATAGTATATCCCAACAGTTCGAGGAGTATATCAGTTCCTTAATC 192  
 QY 2946 ATATAATGTCAACCTTTTACAAACCGGCTCTTATACGCGCAACTTTCGCTCAT 3005  
 DB 193 ATATAATGTCAACCTTTTACAAACCGGCTCTTATACGCGCAACTTTCGCTCAT 252  
 QY 3006 CATTTTGGCCTCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACCTCCGAATC 3065  
 DB 253 CATTTTGGCCTCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACCTCCGAATC 312  
 QY 3066 TGGCTGCATGGCGGCGACACTGTTGAGCTGCTAGTGTGTTGGTGGCTGCCAATGGAT 3125  
 DB 313 TGGCTGCATGGCGGCGACACTGTTGAGCTGCTAGTGTGTTGGTGGCTGCCAATGGAT 372  
 QY 3126 TATTACCTTTATGTC 3141  
 DB 373 TGGTACCTTTATGGGC 388

RESULT 12  
 AV784525/c  
 LOCUS AV784525 360 bp mRNA linear EST 28-MAR-2002  
 DEFINITION AV784525 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-20-N17 3',  
 mRNA sequence.  
 ACCESSION AV784525  
 VERSION AV784525.1 GI:19803315  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 360)  
 AUTHORS Seki.M., Narusaka.M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,  
 Oono.Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,  
 Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.  
 and Shinozaki,K.  
 Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 Unpublished  
 Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msekirc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). This clone is in a  
 modified pBluescript vector as a SstI/XhoI insert. Please visit our  
 web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for  
 further details.

TITLE JOURNAL  
 COMMENT

FEATURES  
 source  
 Location/Qualifiers  
 1..360  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL05-20-N17"  
 /dev\_stage="rosette plants"

/lab\_host="SOLR"  
 /clone\_lib="RAFL5"  
 /note="Site 1: SstI; Site 2: XhoI; subjected to  
 dehydration-treated (1,2,5,10,24 hr)."  
 BASE COUNT 128 a 59 c 69 g 104 t  
 ORIGIN

Query Match 8.3%; Score 360; DB 9; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-47;  
 Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3859 CAGATTTCCTGGAATCGGATTAGATCTTTCAAAGCTAAATGAGTAGAGGAACAGGTT 3918  
 DB 360 CAGATTTCCTGGAATCGGATTAGATCTTTCAAAGCTAAATGAGTAGAGGAACAGGTT 301  
 QY 3919 TTAACCATACAAAGATGTTCCATGGAAGCAAGAAAGTAGAATCTCTCAACAACATC 3978  
 DB 300 TTAACCATACAAAGATGTTCCATGGAAGCAAGAAAGTAGAATCTCTCAACAACATC 241  
 QY 3979 CTATCATTCATGTGGAAACAGAAAGATCCCAACGGATCGGTTGGAACCTCAAGCTTCCA 4038  
 DB 240 CTATCATTCATGTGGAAACAGAAAGATCCCAACGGATCGGTTGGAACCTCAAGCTTCCA 181  
 QY 4039 CATGAGACTCTATTTTCATCTGATCTGTTGTGACTCTGTTTAAAGTTTCAAGACC 4098  
 DB 180 CATGAGACTCTATTTTCATCTGATCTGTTGTGACTCTGTTTAAAGTTTCAAGACC 121  
 QY 4099 ACTGCTACATTTCTTTTCTTTTGGCCCTTTGATTTGTTCTTCTGTCATAGTCTTC 4158  
 DB 120 ACTGCTACATTTCTTTTCTTTTGGCCCTTTGATTTGTTCTTCTGTCATAGTCTTC 61  
 QY 4159 CTGTAACATTTGACTCTGTATTATTCAACAAATCATAACTGTTTAATCTTTTTTTTCC 4218  
 DB 60 CTGTAACATTTGACTCTGTATTATTCAACAAATCATAACTGTTTAATCTTTTTTTTCC 1

RESULT 13  
 AV807633/c  
 LOCUS AV807633 423 bp mRNA linear EST 29-MAR-2002  
 DEFINITION AV807633 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-51-F03 3',  
 mRNA sequence.  
 ACCESSION AV807633  
 VERSION AV807633.1 GI:19841618  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 423)  
 AUTHORS Seki.M., Narusaka.M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,  
 Oono.Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,  
 Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.  
 and Shinozaki,K.  
 Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 Unpublished  
 Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msekirc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially  
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI  
 and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et  
 al., submitted for publication) digested with BamHI and SalI. This  
 clone is in a modified pBluescript vector. Please visit our web  
 site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further  
 details.

TITLE JOURNAL  
 COMMENT

FEATURES  
 source  
 Location/Qualifiers  
 1..423  
 /organism="Arabidopsis thaliana"

BASE COUNT	ORIGIN
143 a	70 c 89 g 121 t

Query Match	8.23;	Score	356.2;	DB	9;	Length	423;
Best Local Similarity	96.08;	Pred. No.	8.3e-47;				
Matches	408;	Conservative	0;	Mismatches	13;	Indels	4;
						Gaps	4;
Qy	3795	GCACCTTGATCTTTAACTTTCACAGCTCAGTTAAACACCCAGTTGATGATCAAGAGGAGGAAGAGA	3854				
Db	423	GCACCTTGATCTTTAACTTTCACAGCTCAGTT-ACACCATTTGATGATCAAGAGGAGGAAGAGA	365				
Qy	3855	AACACAGGATTTCTTTGGAATCGGATTAGATGCTTCAAAGCTAATGAAGTAGAGGAAGAACA	3914				
Db	364	AACACAGGATTTCTTTGGAATCGGATTAGATGCTTCAAAGCTAATGATTAGA-GAAGGACA	306				
Qy	3915	GGTTTTAAACCATACAAAGAGATGTTCCATGGAAGCCAAAGAAAGCTAGAACTCTCAACAAC	3974				
Db	305	GGTTTTAAACCATACAAAGAGATGTTCCAT-GAACCCCAAGAAAGTAGAATCTCTCGACAC	247				
Qy	3975	AATCCTATCATTTGATGGAAACAGAAAGATCCCAAACGGATGCGGTTGGAAACTCAAGCT	4034				
Db	246	CATCCTATCATTTGATGGAAACAGAAAGATCCCAAACGGATGCGGTTGGAAACTCAAGCT	187				
Qy	4035	TCCACATGAGACTCTATTTTTCATCTGATCTGTTCTGTTGTAAGCTGTTTTTAAGTTTTCAA	4094				
Db	186	TCCACATGAGACTCTATTTTTCATCTGATCTGTTCTGTTGTAAGCTGTTTTTAAGTTTTCAA	127				
Qy	4095	GACCACCTGCTACATTTTCTTTTCTTTTGAGGCTTTGTATTTGTTTCTTGTGTCATAGT	4154				
Db	126	GACCACCTGCTACATTTTCTTTTCTTTTGAGGCTTTGTATTTGTTTCTTGTGTCATAGT	67				
Qy	4155	CTTCTCTGAACATTTTGACTCTGTAATTAATCAACAAATCATATAA-CTGTTTAAATCTTTTT	4213				
Db	66	CTTCTCTGAACATTTTGACTCTGTAATTAATCAACCAATCATATAAGCTGTTTAAATCTTTTT	7				
Qy	4214	TTTCC	4218				
Db	6	TTTCC	2				

RESULT 14					
AV827752					
LOCUS	AV827752	580 bp	mRNA	linear	EST 01-APR-2002
DEFINITION	AV827752 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-19-A10 5', mRNA sequence.				

ACCESSION	AV827752
VERSION	AV827752.1
KEYWORDS	GI:19869812
SOURCE	EST.
ORGANISM	Arabidopsis thaliana (thale cress)
	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidops.
REFERENCE	1 (bases 1 to 580)

REFERENCE  
AUTHORS  
1 (bases 1 to 580)  
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
Ono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,  
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.  
and Shinozaki, K.  
large scale analysis of Arabidopsis full-length cdna (2002b)  
TITLE

TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 JOURNAL Unpublished  
 COMMENT Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: mseki@etc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

FEATURES	Location/Qualifiers
source	1. .580
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/clone="RAFL09-19-A10"
	/dev_stage="plants at various developmental stages from germination to mature seeds"
	/lab_host="DH10B"
	/clone_lib="RAFL9"
	/notes="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
BASE COUNT	184 a 82 c 122 g 189 t 3 others

	Query Match	7.8%	Score 339.4;	DB 9;	Length 580;
	Best Local Similarity	82.4%;	Pred. No. 3.7e-44;		
	Matches 463;	Conservative 0;	Mismatches 9;	Indels 90;	Gaps 3;
Qy	1441	TGATATAGACTCGGAAGCCATATACCATACAAGAACAAGTGAAGCTGAAGCTGCACGTGAGGAAG	1500		
Db	108	TTATTAAAGACTCGGAAGCCATATACCATACAAGAACAAGTGAAGCTGCACGTGAGGAAG	167		
Qy	1501	AACATAATAGATTCAATTGAAGCTTTTGAGGCTTTATGGTAGAGCATCGCAGAACATTGAAG	1560		
Db	168	AACATAATAGATTCAATTGAAGCTTTTGAGGCTTTATGGTAGAGCATCGCAGAACATTGA--	225		
Qy	1561	GTTGATTTTTATTTCGCCITTTATATGTCCTATTTTTTTGTGTTGTCAGAGGTTTGTCTTTCAA	1620		
Db	226	-----	225		

Qy	1621	ACTGATTGCTTTTTTTTCATTTTGACAGAACATGTAGCAACAAAAAACTGCTGTCCAGATA	1680
Db	226	-----AGAACATGTAGCAACAAAAAACTGCTGTCCAGATA	259
Qy	1681	AGAAGTCACGCTCAGAAAAATTTTTTCTCCAAGGTAAAAATCGGTAAATTTTGAATATGATGTGTC	1740
Db	260	AGAAGTCACGCTCAGAAAAATTTTTTCTCCAAGGTAAAAATCGGTAAATTTTGAATATGATGTGTC	319

[illegible][illegible]

Qy	1860	GTCTTAAATAGTTCGAGAAATGACAGAGTCAATCGTTTTCTTTTTGTTATGCCCTAT	1919
Dd <td>440 <th>GTCTTAAATAGTTCGAGAAATGACAGAGTCAATCGTTTTCTTTTTGTTATGCCCTAT</th> <td>499</td> </td>	440 <th>GTCTTAAATAGTTCGAGAAATGACAGAGTCAATCGTTTTCTTTTTGTTATGCCCTAT</th> <td>499</td>	GTCTTAAATAGTTCGAGAAATGACAGAGTCAATCGTTTTCTTTTTGTTATGCCCTAT	499

Qy	1920	1979
Db	CTTGTGTTATCCACAACATGTA	CTTGTGTTATCTTTAGTTCTTTAGATTC
	500	558
	CTTGTGTT-ATCCACAACATGTA	CTTGTGTTAGTCTTTAGATTC
		CTTGTGTTAGTCTTTAGATTC

Qy	1980	TTCAGATAAACTTGTGCTTCT	2001
Db	559	TTCAGATAAACTAGCGTCTTCT	580

RESULT 15

RESULT 15

Db 51 TATATGAGATGTTTGTGTTTCATTTTCCAAAAATTTCAAACCTTTGTCCCC 2

Search completed: November 24, 2003, 13:01:50  
Job time : 8562.19 secs

[illegible]

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 18:53:22 ; Search time 15516.7 Seconds  
(without alignments)  
11452.890 Million cell updates/sec

Title: US-10-084-553-1

Perfect score: 4344

Sequence: 1 gcaatgggtcacttaaga.....actcataagtaagcacaaaa 4344

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2889711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sta.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sta.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rtd.\*
- 36: em\_htg\_nam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4344	100.0	4344	6	BD107660	BD107660 Phytochro
2	4344	100.0	5242	8	ATU79156	U79156 Arabidopsis
3	4310	99.2	68415	8	AY050961	AY050961 Arabidops
4	2092.2	48.2	2684	8	BD107661	BD107661 Phytochro
5	1527.6	35.2	2254	6	ATU28422	U28422 Arabidopsis
6	1525.2	35.1	2240	8	AX506805	AX506805 Sequence
7	1349.2	31.1	1827	6	BT001096	BT001096 Arabidops
8	1342	30.9	1612	8	BT001105	BT001105 Arabidops
9	1134	26.1	1134	6	AX509497	AX509497 Sequence
10	1134	26.1	1134	6	AY086711	AY086711 Arabidops
11	306.6	7.1	895	8	AY065156	AY065156 Arabidops
12	305.6	7.0	906	8	AY166504	AY166504 Arabidops
13	223.4	5.1	318	8	AY114583	AY114583 Arabidops
14	168.6	3.9	632	8	ATH530832	ATH530832 Arabidops
15	126	2.9	126	8	ATH530917	ATH530917 Arabidops
16	126	2.9	126	8	ATH531369	ATH531369 Arabidops
17	118.2	2.7	86436	8	AC007323	AC007323 Genomic s
18	118.2	2.7	86436	8	AC013354	AC013354 Genomic s
19	101.6	2.3	95327	8	ATAC010927	ATAC010927 Arabidops
20	100	2.3	96232	8	AB005238	AB005238 Arabidops
21	99.2	2.3	75188	8	A68433	A68433 Sequence 6
22	90.4	2.1	666	6	AP005750	AP005750 Oryza sat
23	88.4	2.0	151517	2	AP004685	AP004685 Oryza sat
24	88.4	2.0	167114	2	PVU420902	PVU420902 Phaseolus
25	85.8	2.0	2172	8	AC126787	AC126787 Medicago
26	85	2.0	124212	8	AC137985	AC137985 Medicago
27	84.2	1.9	131006	2	A68428	A68428 Sequence 1
28	83.2	1.9	2526	6	ATAJ6404	ATAJ6404 Arabidops
29	83.2	1.9	2526	8	AP004460	AP004460 Oryza sat
30	82.8	1.9	148626	8	AX506421	AX506421 Sequence
31	81.6	1.9	1938	6	OSJN00188	OSJN00188 Oryza sat
32	72.8	1.7	186315	2	AC137597	AC137597 Oryza sat
33	72.8	1.7	229796	2	AX660076	AX660076 Sequence
34	72.4	1.7	2160	6	AX660581	AX660581 Sequence
35	70.8	1.6	451	6	AX344550	AX344550 Sequence
36	66.4	1.5	349980	6	AP004255	AP004255 Oryza sat
37	66	1.5	120335	2	AP004071	AP004071 Oryza sat
38	66	1.5	204292	2	AX251342	AX251342 Sequence
39	65.4	1.5	8524	6	AX346743	AX346743 Sequence
40	65.4	1.5	8524	6	AX116988	AX116988 Dictyoste
41	64.4	1.5	331039	3	AY114070	AY114070 Arabidops
42	64.2	1.5	1072	8	AY063952	AY063952 Arabidops
43	64.2	1.5	1429	8	AC117176	AC117176 Dictyoste
44	63.8	1.5	182871	3	AC116367	AC116367 Oryza sat
45	63.4	1.5	134971	2		

ALIGNMENTS

RESULT 1  
BD107660

LOCUS BD107660  
DEFINITION Phytochrome regulated transcription factor for control of higher  
plant development.

ACCESSION BD107660  
VERSION BD107660.1 GI:23202478  
KEYWORDS JP 2002501381-A/1.  
SOURCE unidentified  
ORGANISM unclassified

REFERENCE 1 (bases 1 to 4344)  
AUTHORS Tobin,R., Sun,L. and Wang,Z.Y.  
TITLE Phytochrome regulated transcription factor for control of higher  
plant development

JOURNAL Patent: JP 2002501381-A 1 15-JAN-2002;  
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
COMMENT OS Unidentified  
PN JP 2002501381-A/1  
PD 15-JAN-2002  
PF 17-APR-1998 JP 1998546171  
PR 18-APR-1997 US 08/843572  
PI ROLANINE TOBIN, LIN SUN, ZHI YONG WANG  
PC C12N15/00, C12N15/09, C12N15/29, C12N15/82, A01H1/00, A01H3/00, PC  
A01H5/00  
CC Strandedness: Single;  
CC Topology: Linear;  
CC Phytochrome regulated transcription factor for control of CC  
higher plant  
CC development  
FH Key Location/Qualifiers  
FT source 1.4344  
FT /organism='Unidentified'.  
FEATURES  
source Location/Qualifiers  
1.4344  
/organism='unidentified'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32644'  
BASE COUNT 1303 a 808 c 883 g 1350 t  
ORIGIN  
Query Match 100.0%; Score 4344; DB 6; Length 4344;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGTGGTTCACATTACAAAGACCTGGTCTTCAACACGACAGGTAAACCAATTCCTCTT 60  
DB 1 GCAGTGGTTCACATTACAAAGACCTGGTCTTCAACACGACAGGTAAACCAATTCCTCTT 60  
QY 61 TAACTCTGTGTTGGTTCAGTAACTGAGAACTGGAAGACTCAAAATTCGAGGAAAT 120  
DB 61 TAACTCTGTGTTGGTTCAGTAACTGAGAACTGGAAGACTCAAAATTCGAGGAAAT 120  
QY 121 TGTGTTGTAFTCTTTTCAGGAGGCTTTTGTTCAGAAAGGTCAAGAGCACATACAAAGACA 180  
DB 121 TGTGTTGTAFTCTTTTCAGGAGGCTTTTGTTCAGAAAGGTCAAGAGCACATACAAAGACA 180  
QY 181 TATTAGGAGCAGCTGAATCAAGAGGAGAAAGAGAGAGAGAGGCTTTTTCAGGAGCCAT 240  
DB 181 TATTAGGAGCAGCTGAATCAAGAGGAGAAAGAGAGAGAGAGGCTTTTTCAGGAGCCAT 240  
QY 241 TCATGAATTCGAATGAAGGATATCAAAAGAACTCAACAAAGGCCACGCTCTCTTCA 300  
DB 241 TCATGAATTCGAATGAAGGATATCAAAAGAACTCAACAAAGGCCACGCTCTCTTCA 300  
QY 301 ATCTTTCTCTTTGTAACATAAATTTTCATCCTTTCTCTCTCTCTCTCTCTCTCTCTCT 360  
DB 301 ATCTTTCTCTTTGTAACATAAATTTTCATCCTTTCTCTCTCTCTCTCTCTCTCTCTCT 360  
QY 361 TTTAGTCAAGATATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGACTATATAT 420  
DB 361 TTTAGTCAAGATATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGACTATATAT 420  
QY 421 GAGATGTTTGTTCATTTTCAAAATTTCAAACTTTGTCCTCTCTCTCTCTCTCTCTCTCT 480  
DB 421 GAGATGTTTGTTCATTTTCAAAATTTCAAACTTTGTCCTCTCTCTCTCTCTCTCTCTCT 480  
QY 481 TCATGCATGGTTCAGCTTAATGCTGAATGTTGTAATTAACGATATGGGCTTATGCT 540  
DB 481 TCATGCATGGTTCAGCTTAATGCTGAATGTTGTAATTAACGATATGGGCTTATGCT 540  
QY 541 AAAAGNACAAAACCTTATGGTCTAAAAAATAAGCCCAATATAAAAATATAGGCCAAA 600  
DB 541 AAAAGNACAAAACCTTATGGTCTAAAAAATAAGCCCAATATAAAAATATAGGCCAAA 600  
QY 601 TAAGTTTAGGTCATTTAGAGTGTGAAATAGCGCTGTAGTGAACCGCACAGAGATGCGC 660  
DB 601 TAAGTTTAGGTCATTTAGAGTGTGAAATAGCGCTGTAGTGAACCGCACAGAGATGCGC 660

QY 661 GTTCGATTGTTGTTGAGTACTGCTCTAGATTCCCGGGTCCACTGATGTTTCTAGTGTAT 720  
DB 661 GTTCGATTGTTGTTGAGTACTGCTCTAGATTCCCGGGTCCACTGATGTTTCTAGTGTAT 720  
QY 721 CAGACACGCTGTCGACAAACTGGTGGGAGAGATTAAACCATCTTTAAGTAGGTCCCACTAGAT 780  
DB 721 CAGACACGCTGTCGACAAACTGGTGGGAGAGATTAAACCATCTTTAAGTAGGTCCCACTAGAT 780  
QY 781 CAAGATATTATAACGAATTGACCTTTTAAACCTTTTAAACCTTTTAAACCTTTTAAACCTTT 840  
DB 781 CAAGATATTATAACGAATTGACCTTTTAAACCTTTTAAACCTTTTAAACCTTTTAAACCTTT 840  
QY 841 AGAATACAAAGAGAGGTGTGAAACAAAGTTGATGTTAAGATGGACAAAGATGTAACCTTGAAC 900  
DB 841 AGAATACAAAGAGAGGTGTGAAACAAAGTTGATGTTAAGATGGACAAAGATGTAACCTTGAAC 900  
QY 901 AAAAGCTGAATCATCTCTTCAGCCACTAGTATGTTGACATATGGCAGTCTCTTTTGTAGC 960  
DB 901 AAAAGCTGAATCATCTCTTCAGCCACTAGTATGTTGACATATGGCAGTCTCTTTTGTAGC 960  
QY 961 CTCGAAATAAATAAATAAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTCT 1020  
DB 961 CTCGAAATAAATAAATAAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTCT 1020  
QY 1021 CCAATTTCCGTAGCTCTGCT 1080  
DB 1021 CCAATTTCCGTAGCTCTGCT 1080  
QY 1081 TCTTCT 1140  
DB 1081 TCTTCT 1140  
QY 1141 TGGAACTCTTATCGAAATCCAAGCTGATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1200  
DB 1141 TGGAACTCTTATCGAAATCCAAGCTGATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
QY 1201 TACTTAAGATGATGTTTATGTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260  
DB 1201 TACTTAAGATGATGTTTATGTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260  
QY 1261 TTGTTTTTTGTTTGTGGATTAGTGGAAATTTTCTAAAGAGAACATCTGAAGTTGTGTAGA 1320  
DB 1261 TTGTTTTTTGTTTGTGGATTAGTGGAAATTTTCTAAAGAGAACATCTGAAGTTGTGTAGA 1320  
QY 1321 GGAGCTTAGTATGGAGACAAATTCGCTGGAGAACATCTGGTTATTAAAGGTAATTAAC 1380  
DB 1321 GGAGCTTAGTATGGAGACAAATTCGCTGGAGAACATCTGGTTATTAAAGGTAATTAAC 1380  
QY 1381 TAAATTTTAGGGGAGATGATTGTTTATAGGTCTCAAGATTGAGAAATTTTAAATGAACT 1440  
DB 1381 TAAATTTTAGGGGAGATGATTGTTTATAGGTCTCAAGATTGAGAAATTTTAAATGAACT 1440  
QY 1441 TGATATAGACTCGGAAGCCATATACGATAACAAAGCAACCTGAAAGGTGAGCTGAGGAAG 1500  
DB 1441 TGATATAGACTCGGAAGCCATATACGATAACAAAGCAACCTGAAAGGTGAGCTGAGGAAG 1500  
QY 1501 AACATAATAGATTCATTGAAGCTTTGAGCTTTATGGTAGAGCATGGCAGAAATGGAAG 1560  
DB 1501 AACATAATAGATTCATTGAAGCTTTGAGCTTTATGGTAGAGCATGGCAGAAATGGAAG 1560  
QY 1561 GTTGATTTTATTTCCCTTTATATGCTTATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTT 1620  
DB 1561 GTTGATTTTATTTCCCTTTATATGCTTATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTT 1620  
QY 1621 ACTGATTTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680  
DB 1621 ACTGATTTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680  
QY 1681 AGAAGTCAGCTCAGAAATTTTCTCAGAGTAAATCGGTTAAATTTTGAATATGATGTTTC 1740  
DB 1681 AGAAGTCAGCTCAGAAATTTTCTCAGAGTAAATCGGTTAAATTTTGAATATGATGTTTC 1740

QY	1741	TCATCTTCATTGGCTTAATGCTTAAGACTTAATTGAAGCCAGGCAAGTTTCTGCTCTCT	1800
DB	1741	TCATCTTCATTGGCTTAATGCTTAAGACTTAATTGAAGCCAGGCAAGTTTCTGCTCTCT	1800
QY	1801	TTGCTCTCTTAGTCAGGAGATAGATAGATTACGTTTTTATAGATTTTAGTAATGACCAATAAG	1860
DB	1801	TTGCTCTCTTAGTCAGGAGATAGATAGATTACGTTTTTATAGATTTTAGTAATGACCAATAAG	1860
QY	1861	TCCTAAAAATAGTTGGAGAAATGACGAGATGTAATCGTTTCTTTTGTATGCTATATC	1920
DB	1861	TCCTAAAAATAGTTGGAGAAATGACGAGATGTAATCGTTTCTTTTGTATGCTATATC	1920
QY	1921	TTGCTTAATCCACAAACATGTACATAGATTCTTCAGAGAAATGTTAGTTTCTTTTAGATTCT	1980
DB	1921	TTGCTTAATCCACAAACATGTACATAGATTCTTCAGAGAAATGTTAGTTTCTTTTAGATTCT	1980
QY	1981	TCAGATAAACTTGTGCTCTTCTACCGATTCTGAGGTAGTGCCAAAAAGTGGCTGAGTGCT	2040
DB	1981	TCAGATAAACTTGTGCTCTTCTACCGATTCTGAGGTAGTGCCAAAAAGTGGCTGAGTGCT	2040
QY	2041	AGAAATTTTGAATGTTCTTGTGATAGCAATAGAGGTAAACCATTTTGTATTTTCCAG	2100
DB	2041	AGAAATTTTGAATGTTCTTGTGATAGCAATAGAGGTAAACCATTTTGTATTTTCCAG	2100
QY	2101	TTCTGTCAATTAACCTTGTAGGTGTCATTAGATTCTTTGTTTACGTTTGTAGAG	2160
DB	2101	TTCTGTCAATTAACCTTGTAGGTGTCATTAGATTCTTTGTTTACGTTTGTAGAG	2160
QY	2161	GGTAACAAAACACTCTCTCTCAGGTAGAGAAAGGCTGAAGCTTAAAGGTGTAG	2220
DB	2161	GGTAACAAAACACTCTCTCTCAGGTAGAGAAAGGCTGAAGCTTAAAGGTGTAG	2220
QY	2221	CTATGGGTCAAGCGCTAGACATAGCTATTCCTCTCCAGGCTTAAGCGGTAAACCAACA	2280
DB	2221	CTATGGGTCAAGCGCTAGACATAGCTATTCCTCTCCAGGCTTAAGCGGTAAACCAACA	2280
QY	2281	ATCCTTTATCCTCGAAGACGGGAAGTGGACGATCCTTATGTCAAAAACGGGTGTAATG	2340
DB	2281	ATCCTTTATCCTCGAAGACGGGAAGTGGACGATCCTTATGTCAAAAACGGGTGTAATG	2340
QY	2341	ATGGAAAGAAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGGTGATTTTCATGGTCA	2400
DB	2341	ATGGAAAGAAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGGTGATTTTCATGGTCA	2400
QY	2401	TATGGCATCTTTTTCAGTGTGTCAATGCTCTCTCATGTTTATTAATACAGATTGTGTGC	2460
DB	2401	TATGGCATCTTTTTCAGTGTGTCAATGCTCTCTCATGTTTATTAATACAGATTGTGTGC	2460
QY	2461	TTCTGTTTATAGATGGCCCAATGAAGATCGACAAATCAAGCCTGAAGCAAACTCTGC	2520
DB	2461	TTCTGTTTATAGATGGCCCAATGAAGATCGACAAATCAAGCCTGAAGCAAACTCTGC	2520
QY	2521	AGGAAGACAACCTGTTTCAGATTGTTTTCACCTCATCAGTATCTCTGCTGCATCCTCCATGA	2580
DB	2521	AGGAAGACAACCTGTTTCAGATTGTTTTCACCTCATCAGTATCTCTGCTGCATCCTCCATGA	2580
QY	2581	ATAAAAAGTTGTATAGACATCAAAACGCAAGCACTTTCGCGAGTTCTTGGCTTCAACGGG	2640
DB	2581	ATAAAAAGTTGTATAGACATCAAAACGCAAGCACTTTCGCGAGTTCTTGGCTTCAACGGG	2640
QY	2641	AAGAGTAAAAACAATCTTTTCATCTGCTATTTGAGGTTTTTAAGACGATTAGTACTTTTCA	2700
DB	2641	AAGAGTAAAAACAATCTTTTCATCTGCTATTTGAGGTTTTTAAGACGATTAGTACTTTTCA	2700
QY	2701	TGAAACTAAAAACCGTGGGGAATTAACAGGGAAGTCAGAAATAACAGGGTAAGAAAGGATC	2760
DB	2701	TGAAACTAAAAACCGTGGGGAATTAACAGGGAAGTCAGAAATAACAGGGTAAGAAAGGATC	2760
QY	2761	AAACTCAGATTTGAATGCAAAATCTCTCGAAACCGGTAAATGACCAAGGACCTCAGACTTA	2820
DB	2761	AAACTCAGATTTGAATGCAAAATCTCTCGAAACCGGTAAATGACCAAGGACCTCAGACTTA	2820
QY	2821	TCCGATGCATATCCCTGTGCTAGTGCCATTGGGGAGCTCAATAACAAGTTCTCTATCACA	2880

DB	2821	TCCGATGCATATCCCTGTGCTAGTGCCATTGGGGAGCTCAATAACAAGTTCTCTATCACA	2880
QY	2881	TCCTCTCTCAGAGCCAGATAGTATCCACACAGTTGACAGGATTTATCAGTCGTTTCC	2940
DB	2881	TCCTCTCTCAGAGCCAGATAGTATCCACACAGTTGACAGGATTTATCAGTCGTTTCC	2940
QY	2941	TAAATCATATAATGTCAACCCCTTTTACAAAACACCGGCTCTTTTATATCTGCGCAACTTTCG	3000
DB	2941	TAAATCATATAATGTCAACCCCTTTTACAAAACACCGGCTCTTTTATATCTGCGCAACTTTCG	3000
QY	3001	CTCATCATTTTGGCTCTCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACCTCC	3060
DB	3001	CTCATCATTTTGGCTCTCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACCTCC	3060
QY	3061	GAATCTGGCTGCCATGGCCGCGACCACTGTTGAGCTGCTAGTGTGTTGGTGGCTGCGAA	3120
DB	3061	GAATCTGGCTGCCATGGCCGCGACCACTGTTGAGCTGCTAGTGTGTTGGTGGCTGCGAA	3120
QY	3121	TGGATTATTAACCTTTTATGTGCTCTCTTATAGTTTCAAGTGGTGTTCACCTAGTCTCTCCATC	3180
DB	3121	TGGATTATTAACCTTTTATGTGCTCTCTTATAGTTTCAAGTGGTGTTCACCTAGTCTCTCCATC	3180
QY	3181	TACTTTTGGACCATCATGTGATAGTACACAAAAGCAAGCACTTTTACAAATGTTTC	3240
DB	3181	TACTTTTGGACCATCATGTGATAGTACACAAAAGCAAGCACTTTTACAAATGTTTC	3240
QY	3241	TGTGACAGCCGAGAGCAAGAACACTCCGAGGATCAAAAGGCTCGATCTTCACTGGAATC	3300
DB	3241	TGTGACAGCCGAGAGCAAGAACACTCCGAGGATCAAAAGGCTCGATCTTCACTGGAATC	3300
QY	3301	AGAGGATGTTTGAATAATGAAGTAAACACAGTTTGTCTATGAGCAGCCTTCTGCAACACCTGA	3360
DB	3301	AGAGGATGTTTGAATAATGAAGTAAACACAGTTTGTCTATGAGCAGCCTTCTGCAACACCTGA	3360
QY	3361	GAGTGTGCAAAAGGTTTCAGATGGAGCAGAGACAGAAACAAAGTTGACCGGTCTCTCGT	3420
DB	3361	GAGTGTGCAAAAGGTTTCAGATGGAGCAGAGACAGAAACAAAGTTGACCGGTCTCTCGT	3420
QY	3421	TGCTCAAAACACTTCCGTCGAGTAGTATGATGTTGAGGCGGATGCAATCAAGAGGCAAGA	3480
DB	3421	TGCTCAAAACACTTCCGTCGAGTAGTATGATGTTGAGGCGGATGCAATCAAGAGGCAAGA	3480
QY	3481	GGATGSCACCAATGTTGAGTGAAGAAACGAATGAAGACACTAATAAAACCTCAAACTTC	3540
DB	3481	GGATGSCACCAATGTTGAGTGAAGAAACGAATGAAGACACTAATAAAACCTCAAACTTC	3540
QY	3541	AGAGTCCAATGACCGCCGAGTAGAATCAGCTCCAAATATAACCGATCCATGGGAAGTCTGT	3600
DB	3541	AGAGTCCAATGACCGCCGAGTAGAATCAGCTCCAAATATAACCGATCCATGGGAAGTCTGT	3600
QY	3601	GTCTGACGAGGTACTTTTACTTTGGACTAAAGATCAACTTCTCTTTTATTTTCAATCATTTTCTC	3660
DB	3601	GTCTGACGAGGTACTTTTACTTTGGACTAAAGATCAACTTCTCTTTTATTTTCAATCATTTTCTC	3660
QY	3661	ATATAAATATGTTACATTTCCGGTTCGAATTCCTTCCAAGCTCTCTTCTCCAGAGAGGTAT	3720
DB	3661	ATATAAATATGTTACATTTCCGGTTCGAATTCCTTCCAAGCTCTCTTCTCCAGAGAGGTAT	3720
QY	3721	TGCGCGAAAGTTTACATATCGAAGAAACACAGAGAGGAAGAAACAAACAAGAAC	3780
DB	3721	TGCGCGAAAGTTTACATATCGAAGAAACACAGAGAGGAAGAAACAAACAAGAAC	3780
QY	3781	AAAGATATCCAATGGCAGTTGATCTTAACTTCCAGCTCAGTTTAAACCCAGTTTGTATGATC	3840
DB	3781	AAAGATATCCAATGGCAGTTGATCTTAACTTCCAGCTCAGTTTAAACCCAGTTTGTATGATC	3840
QY	3841	AAAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAAGCTAATGA	3900
DB	3841	AAAGAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAAGCTAATGA	3900
QY	3901	GTAGAGAGAGAACACAGGTTTTTAAACNTACAAAAGATGTTTCCATGGGAAGCCAAAGAAAGTA	3960



Db 3901 GTAGAGGAAGAACAGGTTTTAAACCATACAAAAGATGTTCCATGGAAGCCAAAGAAAGTA 3960  
QY GAATCTCTCAACAAATCTCTATCATTCATGTGGAACAGAAAGATCCAAACCGATCGGT 4020  
Db |||||  
Db 3961 GAATCTCTCAACAAATCTCTATCATTCATGTGGAACAGAAAGATCCAAACCGATCGGT 4020  
QY TGGAACTCAAGCTTCCACATGAGACTCTATTTTCATCTGATCTGTTGTTGTTACTCTGT 4080  
Db |||||  
QY 4021 TGGAACTCAAGCTTCCACATGAGACTCTATTTTCATCTGATCTGTTGTTGTTACTCTGT 4080  
Db |||||  
QY 4081 TTTTAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTCTTTGAGGCTTGTGTTGTT 4140  
Db |||||  
QY 4081 TTTTAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTCTTTGAGGCTTGTGTTGTT 4140  
QY 4141 TCTTGTCTCATAGTCTTCTGTAACTTTGACTCTGTATTTATTTCAACAAATCATAAATCTG 4200  
Db |||||  
QY 4141 TCTTGTCTCATAGTCTTCTGTAACTTTGACTCTGTATTTATTTCAACAAATCATAAATCTG 4200  
QY 4201 TTTAATCTTTTCTTCCACCTGGAAGAACTTCTACTCAAGGGCTTGTGTTCTTGATA 4260  
Db |||||  
QY 4201 TTTAATCTTTTCTTCCACCTGGAAGAACTTCTACTCAAGGGCTTGTGTTCTTGATA 4260  
QY 4261 TATGCAACGACAGAGTCTCAAAACGTAATCTTTAGCCCATCCATCAACCTTAAGTTGCT 4320  
Db |||||  
QY 4261 TATGCAACGACAGAGTCTCAAAACGTAATCTTTAGCCCATCCATCAACCTTAAGTTGCT 4320  
QY 4321 CATACTCATAGTAAGCAACAAA 4344  
Db |||||  
QY 4321 CATACTCATAGTAAGCAACAAA 4344

RESULT 2  
ATU79156  
LOCUS Arabidopsis thaliana Myb-related transcription factor CCA1 gene, complete cds. 5242 bp DNA linear PLN 27-SEP-2001  
DEFINITION U79156  
ACCESSION U79156.1 GI:4090568  
VERSION  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM  
REFERENCE Wang,Z.-Y., Kenigsbuch,D., Sun,L., Harel,E., Ong,M.S. and Tobin,E.M. 9144958  
AUTHORS A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene Plant Cell 9 (4), 491-507 (1997)  
JOURNAL 97290202  
MEDLINE  
REFERENCE Wang,Z.-Y., Kenigsbuch,D., Sun,L., Harel,E., Ong,M.S. and Tobin,E.M. 9144958  
AUTHORS Direct Submission  
TITLE Submitted (21-NOV-1996) MCD Biology, University of California at Los Angeles, 405 Hilgard, Los Angeles, CA 90095, USA  
JOURNAL Location/Qualifiers  
FEATURES  
1..5242  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="2"  
join(1909..2097,2181..2268,2347..2458,2547..2608,3088..3282,3370..3543,3627..4508,4579..5118)  
1909..2229  
1909..2097  
/number=1  
2098..2180  
/number=1  
2181..2268  
/number=2

mRNA  
5'UTR  
exon  
intron  
exon

CDS  
join(2230..2268,2347..2458,2547..2608,3088..3282,3370..3543,3627..4508,4579..4941)  
/note="Myb-related transcription factor"  
/codon\_start=1  
/product="CCA1"  
/protein\_id="AAC98813.1"  
/db\_xref="GI:4090569"  
/translation="METNSSGDELVIKTRPKYTIKQERWTEEEHNRFEALRLYCR  
AMQTEEHVATHTAVQIRSHAQKFFSKVEKEAEKGVAMQALDIAIPPRPKRKN  
PYPRKTSSTILMSKTGVNDKESLGSEKVSHPMANEDRQOSKPEKTLQEDNCSDC  
FTHQYLSAASSMKNKSCIEITSNASTREFLPSREEGSONNRVKESNDLNSLENG  
EQGPOTYPHIEVLVPLGSSITSSLSHPSEDPDHPHTVAGDYQSPFNHIMSTLQTP  
ALYTAATFASSEWPPDSSGSPVPGNSPPLAAMAAATVAASAWAANGLLLPLCAPL  
SSGPTSHPPSTFGSCDVEYTKASTLOHGSVQSRQEHSEASKBSLSDSEVENKS  
KVCHEQPSATPESDAKSGDAGDRQVDRSSCGSNTPSSSDVDEADASERQBDGNG  
EVEKNEDTNKQTSNARRKISNITDPKWSVDEGRIFAPQALFSEVLVPSFTY  
REHEEEQQOQOERYPMALDINFTAQLTTPVDQOEKNTGFLGIGLDASKLMSRGT  
GFKPYKRCMEAKESRIILNNPPIIHVEQDKPKMRLETOAST"  
2269..2346  
/number=2  
2347..2458  
/number=3  
2459..2546  
/number=3  
2547..2608  
/number=4  
2609..3087  
/number=4  
3088..3282  
/number=5  
3283..3369  
/number=5  
3370..3543  
/number=6  
3544..3626  
/number=6  
3627..4508  
/number=7  
4509..4578  
/number=7  
4579..5118  
/number=8  
4942..5118  
BASE COUNT 1514 a 982 c 1067 g 1679 t  
ORIGIN  
Query Match 100.0%; Score 4344; DB 8; Length 5242;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGTGGTTCACTTACAAGAACCTGGTCTTCAAAACCAGACAGAGTTAAACCAATTTCTCTT 60  
Db |||||  
QY 899 GCAGTGGTTCACTTACAAGAACCTGGTCTTCAAAACCAGACAGAGTTAAACCAATTTCTCTT 958  
QY 61 TAATCTCTGTTGTTGGTTCATGTAATCTAGAGATGGAAGACTCAAAATTTCTGAGGAAT 120  
Db |||||  
QY 959 TAATCTCTGTTGTTGGTTCATGTAATCTAGAGATGGAAGACTCAAAATTTCTGAGGAAT 1018  
QY 121 TGTGTTTATCTGTTTCAGGAGGCTTTGTTTGAAGAGTCAAGAGCATACAAAGACA 180  
Db |||||  
QY 1019 TGTGTTTATCTGTTTCAGGAGGCTTTGTTTGAAGAGTCAAGAGCATACAAAGACA 1078  
QY 181 TATTAGGAGCAGCTGAATCAAGAGGAAGAAAGAGAGAGCCCTTTTGGGCCAT 240  
Db |||||  
QY 1079 TATTAGGAGCAGCTGAATCAAGAGGAAGAAAGAGAGAGCCCTTTTGGGCCAT 1138  
QY 241 TCATGAATTTGAATGAAGGATATCAAAAGAAATCAACAAAGGCCAGCTCTCTCTTCA 300  
Db |||||  
QY 1139 TCATGAATTTGAATGAAGGATATCAAAAGAAATCAACAAAGGCCAGCTCTCTCTTCA 1198  
QY 301 ATCTTCTCTTCTGTAATAATAATTTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
Db |||||  
QY 1199 ATCTTCTCTTCTGTAATAATAATTTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1258

QY	361	TTTAGCTCAAGATATCATCCATTTATGTCAAAGTGTGTAAATTCCTCAAGACTATAT	420
Db	1259	TTTAGCTCAAGATATCATCCATTTATGTCAAAGTGTGTAAATTCCTCAAGACTATAT	1318
QY	421	GAGATGTTTTGTTTCAATTTTCCAAAATTTCAAACCTTTGTCGCCATTTAGTCTTACCT	480
Db	1319	GAGATGTTTTGTTTCAATTTTCCAAAATTTCAAACCTTTGTCGCCATTTAGTCTTACCT	1378
QY	481	TCATGCATGTTAGCTTACGTTAACTGCTGAATGTTTGAATACGATATGGGCTTATGCT	540
Db	1379	TCATGCATGTTAGCTTACGTTAACTGCTGAATGTTTGAATACGATATGGGCTTATGCT	1438
QY	541	AAAAAGAACAAAACCTTTATGGGCTTAAAAAAATAAGCCCAATATAAAATCTATGGCCAAA	600
Db	1439	AAAAAGAACAAAACCTTTATGGGCTTAAAAAAATAAGCCCAATATAAAATCTATGGCCAAA	1498
QY	601	TAAGTTTAGGTCATTTAGAGTGTGAGAAATAGCGGTGTAGTGAACCGACGAGAAATGCGC	660
Db	1499	TAAGTTTAGGTCATTTAGAGTGTGAGAAATAGCGGTGTAGTGAACCGACGAGAAATGCGC	1558
QY	661	GTTCCGATGTTGTTGAAGTAGTCTAGATTCCTGGGTCCACTGATGTTTCTAGTGAT	720
Db	1559	GTTCCGATGTTGTTGAAGTAGTCTAGATTCCTGGGTCCACTGATGTTTCTAGTGAT	1618
QY	721	CAGACACGTGTGACAAAACCTGGTGGGAGAGATTAAACGATCTTAAAGTAGTCCCACTAGAT	780
Db	1619	CAGACACGTGTGACAAAACCTGGTGGGAGAGATTAAACGATCTTAAAGTAGTCCCACTAGAT	1678
QY	781	CAAGATATTATTAACGAAATGACCTTTTAAACCTTTCAAGGTAGTCCCGAACTCGTGGCT	840
Db	1679	CAAGATATTATTAACGAAATGACCTTTTAAACCTTTCAAGGTAGTCCCGAACTCGTGGCT	1738
QY	841	AGAAATCAAGAGAGGTTGTGAACAAGTTCATGTTTAAAGTAGGACAGAAATGTAACCTGAAC	900
Db	1739	AGAAATCAAGAGAGGTTGTGAACAAGTTCATGTTTAAAGTAGGACAGAAATGTAACCTGAAC	1798
QY	901	AAAAAGTGAATCATCTCTTCAGCCACTAGTATGTTGACATATGCGCATGTTTCTTTGTAGC	960
Db	1799	AAAAAGTGAATCATCTCTTCAGCCACTAGTATGTTGACATATGCGCATGTTTCTTTGTAGC	1858
QY	961	CTCGAAATAAATTAATAAGTTTGAGGTTTAAAGATAATTATAGTGGCTGAGATTTCT	1020
Db	1859	CTCGAAATAAATTAATAAGTTTGAGGTTTAAAGATAATTATAGTGGCTGAGATTTCT	1918
QY	1021	CCATTTCCGTAGCTTCGTCTCTTTCTTTGTTTCAATGATCAAGCAAAATCACTTCT	1080
Db	1919	CCATTTCCGTAGCTTCGTCTCTTTCTTTGTTTCAATGATCAAGCAAAATCACTTCT	1978
QY	1081	TCCTCTCTCTCTCTCGATTTCTTACTGTTTCTTATCAACGAAATCTGGAATTAATAA	1140
Db	1979	TCCTCTCTCTCTCTCGATTTCTTACTGTTTCTTATCAACGAAATCTGGAATTAATAA	2038
QY	1141	TGGAATCTTTATCGAAATCCAAGCTGATTTGTTTCTTTTCAATGAAATCATCTCTTAAAGG	1200
Db	2039	TGGAATCTTTATCGAAATCCAAGCTGATTTGTTTCTTTTCAATGAAATCATCTCTTAAAGG	2098
QY	1201	TACTTAAGATGATTTATTTGTCATGCTCTTCTTATTTGTTGATGAATAACTTGACTTGA	1260
Db	2099	TACTTAAGATGATTTATTTGTCATGCTCTTCTTATTTGTTGATGAATAACTTGACTTGA	2158
QY	1261	TTGTTTTTTGTTTGTGGATTTAGTGGAAATTTTGTAAAGAGAGATCTGAAATTTGTGTAGA	1320
Db	2159	TTGTTTTTTGTTTGTGGATTTAGTGGAAATTTTGTAAAGAGAGATCTGAAATTTGTGTAGA	2218
QY	1321	GGAGCTTAGTGTAGGAGACAAATTCGTCTGGAGAGATCTGGTTATTAAAGGTAATAAAC	1380
Db	2219	GGAGCTTAGTGTAGGAGACAAATTCGTCTGGAGAGATCTGGTTATTAAAGGTAATAAAC	2278
QY	1381	TAAATTTTAAAGGGGAAGATGATTTGTTTAAAGGTGTCAAAGATTTGAGAACTTAACT	1440
Db	2279	TAAATTTTAAAGGGGAAGATGATTTGTTTAAAGGTGTCAAAGATTTGAGAACTTAACT	2338

QY	1441	TGATATAGACTCGAAGCCATATACGATAACAAAGCAACGCTGAAAGGTGACTGAGAAG	1500
Db	2339	TGATATAGACTCGAAGCCATATACGATAACAAAGCAACGCTGAAAGGTGACTGAGAAG	2398
QY	1501	AACATAATAGATTCATTGAAGCTTTGAGGCTTTATGCTAGAGCATGGCAGAAATTTGAAG	1560
Db	2399	AACATAATAGATTCATTGAAGCTTTGAGGCTTTATGCTAGAGCATGGCAGAAATTTGAAG	2458
QY	1561	GTTGATTTTTTATTTCCCTTTATATGCTTATTTTGTGTTTGCAGAGGTTTCTCTTCAA	1620
Db	2459	GTTGATTTTTTATTTCCCTTTATATGCTTATTTTGTGTTTGCAGAGGTTTCTCTTCAA	2518
QY	1621	ACTGATTTCTCTTTTTCATTTTGGACAGAACATGTAGCAACAAAACTGCTGTCCAGATA	1680
Db	2519	ACTGATTTCTCTTTTTCATTTTGGACAGAACATGTAGCAACAAAACTGCTGTCCAGATA	2578
QY	1681	AGAAGTCACGCTCAGAAATTTTTCCTCAAGGTAAAAATCGGTTAAATTTTGAATGATGTTT	1740
Db	2579	AGAAGTCACGCTCAGAAATTTTTCCTCAAGGTAAAAATCGGTTAAATTTTGAATGATGTTT	2638
QY	1741	TCATCTTCAATTCGCTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTT	1800
Db	2639	TCATCTTCAATTCGCTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTT	2698
QY	1801	TTGCTTTCTTAGTCAGGAGATAGATAGATTAACGTTTAAAGCTTTTAAAGCTTTTAAAGCTTT	1860
Db	2699	TTGCTTTCTTAGTCAGGAGATAGATAGATTAACGTTTAAAGCTTTTAAAGCTTTTAAAGCTTT	2758
QY	1861	TCTTAAATAGTTGGAGAAATGACGAGATGTAATCGTTTCTTTTCTTTTATGCTATATC	1920
Db	2759	TCTTAAATAGTTGGAGAAATGACGAGATGTAATCGTTTCTTTTCTTTTATGCTATATC	2818
QY	1921	TTGTTTAAATCCAAACATGTACATAGATTTCTCAGAAAGATGTTAGTCTTTTAGATTCT	1980
Db	2819	TTGTTTAAATCCAAACATGTACATAGATTTCTCAGAAAGATGTTAGTCTTTTAGATTCT	2878
QY	1981	TCAGATAAACTTGTGCTTCTTACCGATTCGAGGTAGTGGCAAAAGTGGGCTGAGTCT	2040
Db	2879	TCAGATAAACTTGTGCTTCTTACCGATTCGAGGTAGTGGCAAAAGTGGGCTGAGTCT	2938
QY	2041	AGAAATTTTGAATGTTCTTGTGATAGCCATAGAGGTAAACCAATTTTGTATTTTCCAG	2100
Db	2939	AGAAATTTTGAATGTTCTTGTGATAGCCATAGAGGTAAACCAATTTTGTATTTTCCAG	2998
QY	2101	TTCTGTCATTTAAACTTGTAGTGTCTTACATTTTGTGTTTGTGTTTGTGTTTGTAGAG	2160
Db	2999	TTCTGTCATTTAAACTTGTAGTGTCTTACATTTTGTGTTTGTGTTTGTGTTTGTAGAG	3058
QY	2161	GGTAAACAAACTTACTCTCATCTCTCAGGTAGAGAAAGAGGCTGAAGCTTAAAGGTGTAG	2220
Db	3059	GGTAAACAAACTTACTCTCATCTCTCAGGTAGAGAAAGAGGCTGAAGCTTAAAGGTGTAG	3118
QY	2221	CTATGGGTCAAGCGCTAGACATAGCTATTTCTCTCCAGGCTTAAAGCTTAAACCAACA	2280
Db	3119	CTATGGGTCAAGCGCTAGACATAGCTATTTCTCTCCAGGCTTAAAGCTTAAACCAACA	3178
QY	2281	ATCCTTATCTCGAAAGACGGAAAGTGGAAACGATCCTTATGTCAAAAAACGGGTGTAATG	2340
Db	3179	ATCCTTATCTCGAAAGACGGAAAGTGGAAACGATCCTTATGTCAAAAAACGGGTGTAATG	3238
QY	2341	ATGGAAGAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGGTGATTTTCATGCTCA	2400
Db	3239	ATGGAAGAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGGTGATTTTCATGCTCA	3298
QY	2401	TATGCACTCTTTTTCAGTGTGTACATTTGCTCTCATGTTTATTAATACAGATTTGTGTC	2460
Db	3299	TATGCACTCTTTTTCAGTGTGTACATTTGCTCTCATGTTTATTAATACAGATTTGTGTC	3358
QY	2461	TTGCTTTTATAGATGGCCATGAGATCGAACCAATCAAGGCTTGAAGAGAAAACTCTGTC	2520
Db	3359	TTGCTTTTATAGATGGCCATGAGATCGAACCAATCAAGGCTTGAAGAGAAAACTCTGTC	3418
QY	2521	AGGAAGACAACTGTTTTCAGATTTGTTTCACTCATCAGTATCTCTCTGCTGCTCATCTCA	2580



```

REFERENCE 3 (bases 1 to 68415)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@igr.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:6598455.
FEATURES
     source             1..68415
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /cultivar="Columbia"
                        /db_xref="taxon:3702"
                        /chromosome="2"
                        /map="C1C06C03"
                        /clone="F19D11"
     misc_feature       1..516
                        /notes="overlap with BAC clone T3A4
                        (AC005819:55355..55870)."
     gene               69..2125
                        /genes="At2g46730"
     mRNA              join(<69..118,1498..1713,1897..2125)
                        /notes="synonym: F19D11.1; predicted by gscan"
                        /genes="At2g46730"
     CDS                join(69..118,1498..1713,1897..2125)
                        /gene="At2g46730"
                        /codon_start=1
                        /product="unknown protein"
                        /protein_id="AAC33492.2"
                        /db_xref="GI:20197316"
                        /translations="MGIAELPRLCLCLNRHIDGTNESNVVVKLRNASDLIWEVK
MEGRRLIQGLKLPTRSGVDVYVFRDGLLFFVTFGSPSCYIRLPLPGILMADGVE
KRTITLVIRYGVKREMSLAKPDYRGMRSSSDTSLWSIDLKGVNGAVETCGVTFHR
LDVY"
     gene               3626..4289
                        /gene="At2g46735"
                        /notes="synonym: F19D11.26; supported by full length cDNA:
Ceres: 105240"
     mRNA              3626..4289
     CDS                3684..4166
                        /gene="At2g46735"
                        /codon_start=1
                        /product="Expressed protein"
                        /protein_id="AA015019.1"
                        /db_xref="GI:20197318"
                        /translations="MMTVKPAAMPVQFHMRTVATSVAYSQPPPHNVGDIHRRFINVP
TKSQTSEVGLGQOILLRTVDNPRKPIPKETIEISGDVLMWAIQRTAQKRTNAGK
KTKTKNRLSGVELSSAGEPTGDNGVDYSNVTPLRIKSDWGHRLLEEFEKLLKEFQNTLE
"
     gene               complement(4391..6664)
                        /gene="At2g46740"
                        /notes="synonym: F19D11.2"
     mRNA              complement(join(<4391..5103,5201..5302,5561..5816,
5963..6664))
     CDS                /gene="At2g46740"
                        complement(join(4391..5103,5201..5302,5561..5816,
5963..6664))
                        /gene="At2g46740"
                        /codon_start=1
                        /product="unknown protein"
                        /protein_id="AAC33493.1"
                        /db_xref="GI:3510249"
                        /translations="MAFGYSPSYCSFWRTLGLLYCLFTLVHTVISTPDPVKCVSGN
TNCIVTNSGAFDRCTCAANVAYPTTEAELVSVAAATKAGRKMRVTRVSHSPK
LTCTDNGDLFTSTKFLNHTVTDATMTLVESGVTLRQLIAEAAKVGALPYAVW
WGVTVCGMMGTGAHSSLMGKGSVAHDYVTEIRMVSPGSVNDGFAKIRVLSETTYNE
ENAAKVLGVLSOVTPELOPMFKESLKYNRNDLDENDEALTGKKHFAFVWL
PSGKVYVRMDRVAVNTLGNLYDFPFRSQSAVLAITRSEETQETLRDAHGCV
TATTISTLFTSYGNTNGITFTGYVIGSQNRMMSSGCLDLEDKLTSSACAMDSR
KETEADFDITYRANDPLTPRLYEDFIEIEQIALFKYNALPHWGNRNLAFDGVIK
KYKNAPAFVKESYDPNGLFSSEWTDQILGKNGVTIIVKDGCALEGLCVCSGD
PSKGYLCRPGKYKEARVCTHPVK"

```

```

repeat_region
gene
/pt_family="AT_rich"
complement(7286..7309)
complement(7652..10055)
/genes="At2g46750"
/notes="synonym: F19D11.3"
complement(join(<7652..8379,8469..8570,9069..9324,
9429..10055))
/genes="At2g46750"
complement(join(7652..8379,8469..8570,9069..9324,
9429..10055))
/genes="At2g46750"
/product="unknown protein"
/protein_id="AAC33494.1"
/db_xref="GI:3510250"
/translations="MHTAVSTPPDDPVKCVSGNCTVNTSYGAFDRSTCAANVAV
PKTEAELVSVAAATGAGRKMRVTRVSHSITKLVCCTDGTGEGLFISTKFLNHTVQADA
TAMITVESGTLRQLIAEAAKVGALPYAPYVWGLTVGGMMGTGAHSSLMGKGSVA
HDYVTEIRMVSPGSVNEGFAKIRILSETTYNEFNAAKVSLGVLSOVTPELOPMF
KRSLTITMRNDSDFEQAVTFGKKHFAFVWLPVSGKVYVRMDRVAVNTLGNLYDF
FLPFRSOLSAAIIRTSETOFRFRDANGKCVGATIISSFTLPAPSVGLTNGIIFTC
YPVVGSQNRMMSSGCLDLSQDGLITACANDSRIGKEFFHOTTLSVLPVTKVKSFLSDI
KSLVKIEQKSLCGLLEHYGLMRYVTSPPALYKETEALDFDITYYAKADPLTPRLYE
DFIEIEQIALFKYNALPHWGNRNLAFDGVIRKYNNAFLKVKDSYDPKGLFSSEW
TDQILGKNGASIVKDGCALEGLCISKDAHCAPAKGYLCRPGKYKEARVCTHRDGI
ISVI"
10960..11002
/pt_family="AT_rich"
11813..11836
/pt_family="(A)n"
complement(12237..14375)
/genes="At2g46760"
/notes="synonym: F19D11.4"
complement(join(<12237..12976,13089..13190,13320..13575,
13662..14375))
/genes="At2g46760"
complement(join(12237..12976,13089..13190,13320..13575,
13662..14375))
/genes="At2g46760"
/codon_start=1
/product="unknown protein"
/protein_id="AAC33495.1"
/db_xref="GI:3510251"
/translations="MAFTSPSPSYCSLNAAFWRTIFVHVCISITLVFTTISTPDPVKVC
VSNCTVNTSYGAFDRSTCAANVAYPTTEAELVSVAAATKAGRKMRVTRVSH
SITKLACTDGTGDLISTKFLNHTVTDATMTLVESGVTLRQLIAEAAKVGALPY
APIWGLTVGGMMGTGAHSSLMGKGSVAHDYVTEIRIVSPGSVNDGFAKVRVLRFT
FVWLPQSGKVYVRMDRVAVNTLGNLYDFPFRSOLSAALAIIRSEETQERFRDAN
GKAGATLITSTLFATSYGLTNGMIFTGYVIGSQNRMMSSGCLDLSQDGLITACP
WDSRIKSEFFHOTTFSIPLTVKSFINDIKSLVKIESKLCVLELYGLMRYVTSPP
AVLKGTEALDFDITYYAKADPLTPRLYEDFIEIEQIALFKYNALPHWGNRNLAED
GVTKYKNVPAFLVKESYDPNGLFSSEWTDQILGKNGVTIIVKDGCALEGLCVCSGD
AHCAPTGYCFRPGKYKEARVCTHRADDISVIOSLSY"
complement(14746..14769)
/pt_family="(TA)n"
14817..14839
/pt_family="AT_rich"
15537..15556
/pt_family="(A)n"
15764..15791
/pt_family="AT_rich"
15978..16009
/pt_family="AT_rich"
complement(16139..16202)
/pt_family="(TA)n"
complement(18735..18755)
/pt_family="AT_rich"
complement(19052..19082)
/pt_family="AT_rich"
complement(19096..19140)
/pt_family="AT_rich"
complement(19507..22028)

```

/gene="At2g46770"  
/note="synonym: F19D11.5"  
complement(join(<19507. .19622,20137. .20749,21247. .21533,  
21845. .>22028))  
/gene="At2g46770"  
complement(join(19507. .19622,20137. .20749,21247. .21533,  
21845. .22028))  
/gene="At2g46770"  
/codon\_start=1  
/product="NAM (no apical meristem)-like protein"  
/protein\_id="AAC3356.1"  
/db\_xref="GI:3510262"  
/translation="MSKMSISVNGSQVPPGFRPHPTBELLQYLRKKNVIEIDL  
DVIRDLNKLDPWDQEMCKIGTTPQNDWFFSHKDKYPTGTRTNRAAAGFWKAT  
GRKIIYSNGRRICMRKTLVYKGRAPHGOKSDWHEYRLDDNIIISPEDVTVHEVVS  
IIGEAODEGVVCRIFPKKNLHKTLNPSVGGASLSGGGDPKTTSSQIJNEDTLDOF  
LELMGRSCKEELNDPMLPNLESNSOAINNHCHVSSPDTNHNHVSNNVDSFVTS  
WAALDRIVASQLNGPSYSITAVNESHVGHDLALPSRSPYPSLNKRSASYHAGLTQE  
YTEMELWNITSSLSPPGFCHVSNVLLVLLRLQLQFQFPQRPQRFHDLSSP  
QMQLSLH"  
21019. .21070  
/repeat\_region  
/rpt\_family="AT\_rich"  
complement(23224. .23252)  
/repeat\_region  
/rpt\_family="AT\_rich"  
24240. .24279  
/repeat\_region  
/rpt\_family="AT\_rich"  
25209. .25229  
/repeat\_region  
/rpt\_family="AT\_rich"  
complement(26557. .26579)  
/repeat\_region  
/rpt\_family="AT\_rich"  
complement(26928. .26979)  
/repeat\_region  
/rpt\_family="AT\_rich"  
28087. .28145  
/repeat\_region  
/rpt\_family="(TAAAA)n"

Query Match 99.28; Score 4310; DB 8; Length 68415;  
Best Local Similarity 99.98; Pred. No. 0;  
Matches 4343; Conservative 0; Mismatches 0; Indels 3; Gaps 3;  
QY 1 GCAGTGGTTCACITACAAAGAACCTGGTCTTCAACACGACAGAGGTTAACCAATCTCTCTT 60  
Db 43882 GCAGTGGTTCACITACAAAGAACCTGGTCTTCAACACGACAGAGGTTAACCAATCTCTCTT 43941  
QY 61 TAACCTCTGTTGGTGGTGCATGTAATCTAGAGATGGAAGACTCAAAATCTCGAGGAAT 120  
Db 43942 TAACCTCTGTTGGTGGTGCATGTAATCTAGAGATGGAAGACTCAAAATCTCGAGGAAT 44001  
QY 121 TGTGTTGTTATCTGTTTCAGGAGGCTTTGTTTCAGAGAGGTTCAAGACACATACAAAGACA 180  
Db 44002 TGTGTTGTTATCTGTTTCAGGAGGCTTTGTTTCAGAGAGGTTCAAGACACATACAAAGACA 44061  
QY 181 TATTAGGAGCAGCTGAATCAAAAGGAGGAAGAAGAAGAGGCTTTTTCAGGAGCAT 240  
Db 44062 TATTAGGAGCAGCTGAATCAAAAGGAGGAAGAAGAAGAGGCTTTTTCAGGAGCAT 44121  
QY 241 TCATGAATTCGAATGAAGGATATCAAAAGAACTCAACACAAAGGCCACGCTTCCTTCCTCA 300  
Db 44122 TCATGAATTCGAATGAAGGATATCAAAAGAACTCAACACAAAGGCCACGCTTCCTTCCTCA 44181  
QY 301 ATCTTTCTCTTCTGTTGTAACATAAATTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360  
Db 44182 ATCTTTCTCTTCTGTTGTAACATAAATTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 44241  
QY 361 TTTAGCTCAAAGTATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGACTATATAT 420  
Db 44242 TTTAGCTCAAAGTATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGACTATATAT 44301  
QY 421 GAGATGTTTGTTCATTTCCAAAATTTCCAACTTTGTCCTCCCATTTAGTCTTACCCT 480  
Db 44302 GAGATGTTTGTTCATTTTCCAAAATTTCCAACTTTGTCCTCCCATTTAGTCTTACCCT 44361  
QY 481 TCATGATGGTGTAGCTTAACTGCTGAATGTTGTAATTAAGCATATGGCCTTATGCT 540

Db 44362 TCATGATGGTGTAGCTTAACTGCTGAACCTGTTGAATAACGATATATGGCCTTATGCT 44421  
QY 541 AAAAGAACAAAACCTTATGGGTCTAAAAAAAATAAGCCCAATATAAAACTATATGGCCCAA 600  
Db 44422 AAAAGAACAAAACCTTATGGGTCTAAAAAAAATAAGCCCAATATAAAACTATATGGCCCAA 44481  
QY 601 TAAATTTAGCTGCATTTAGAGTGTGAGAAATAGCGGCTGTAGTGAACCCGACGAGAAATCGGC 660  
Db 44482 TAAATTTAGCTGCATTTAGAGTGTGAGAAATAGCGGCTGTAGTGAACCCGACGAGAAATCGGC 44541  
QY 661 GTTCGATTTGTTGGTGAAGTAGTCTGTAGATTCCCGGGTCCACTGATGTTTCTTAGTGTAT 720  
Db 44542 GTTCGATTTGTTGGTGAAGTAGTCTGTAGATTCCCGGGTCCACTGATGTTTCTTAGTGTAT 44601  
QY 721 CAGACAGCTCTCCACAAACTGGTGGGAGAGATTAACCATCTTAAGTAGGTCCTCCTAGTAT 780  
Db 44602 CAGACAGCTCTCCACAAACTGGTGGGAGAGATTAACCATCTTAAGTAGGTCCTCCTAGTAT 44661  
QY 781 CAAGATATTATAACGAATTCACCTTTTAAACCTTTTACGTTAGTCCCGGAACCTCGTGGCCT 840  
Db 44662 CAAGATATTATAACGAATTCACCTTTTAAACCTTTTACGTTAGTCCCGGAACCTCGTGGCCT 44721  
QY 841 AGAATCAAAAGAGGTTGTGAAACAAGTTGATGTTAAGATCGACAAAGAAATGTAACCTGAAC 900  
Db 44722 AGAATCAAAAGAGGTTGTGAAACAAGTTGATGTTAAGATCGACAAAGAAATGTAACCTGAAC 44781  
QY 901 AAAAGCTGAATCATCTCTTCAGGCACCTAGTATGTTGACATATGGCAGTGTCTTTTGTAGC 960  
Db 44782 AAAAGCTGAATCATCTCTTCAGGCACCTAGTATGTTGACATATGGCAGTGTCTTTTGTAGC 44841  
QY 961 CTCGAAATAAATAAATTTAAAAAGTTTGAGCTTTAAAGATAAATATATAGTGGCTGAGATTCT 1020  
Db 44842 CTCGAAATAAATAAATTTAAAAAGTTTGAGCTTTAAAGATAAATATATAGTGGCTGAGATTCT 44901  
QY 1021 CCATTTCCGAGCTTCGCTCTCTTTTCTTTTTCATTCATCAAAAGCAAAATCACTTCT 1080  
Db 44902 CCATTTCCGAGCTTCGCTCTCTTTTCTTTTTCATTCATCAAAAGCAAAATCACTTCT 44961  
QY 1081 TCTTCTCTCTCTCTCGATTTCTTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140  
Db 44962 TCTTCTCTCTCTCTCGATTTCTTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 45021  
QY 1141 TGGAACTTTTATCGAATCCAAAGCTGATTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
Db 45022 TGGAACTTTTATCGAATCCAAAGCTGATTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 45081  
QY 1201 TACTTAAGATGTAATTTATGTCATGCTCTTTCTTATTTGTTTGTATGAATAAATCTGACTTGA 1260  
Db 45082 TACTTAAGATGTAATTTATGTCATGCTCTTTCTTATTTGTTTGTATGAATAAATCTGACTTGA 45141  
QY 1261 TTCTTTTTTGTGTTTGTGGAATTTCTGTAAGAGAGATCTGAAGTTGTGTAGA 1320  
Db 45142 TTGTTTTTTGTTTGTGGAATTTCTGTAAGAGAGATCTGAAGTTGTGTAGA 45201  
QY 1321 GGAGCTTAGTGTGAGACAAATTCGTCGGAAGATCTGGTTATTAAAGTAAAT - AA 1379  
Db 45202 GGAGCTTAGTGTGAGACAAATTCGTCGGAAGATCTGGTTATTAAAGTAAAT - AA 45261  
QY 1380 CTAATTTTGGGGGAAGATGATGTTTGTAGTGTCAAAGATTGAGAAATTTTAAATGAAC 1439  
Db 45262 CTAATTTTGGGGGAAGATGATGTTTGTAGTGTCAAAGATTGAGAAATTTTAAATGAAC 45321  
QY 1440 TTGATATAGACTCGGAAGCCATATACGATAACAAACACGCTGAAGGTGACTGAGAA 1499  
Db 45322 TTGATATAGACTCGGAAGCCATATACGATAACAAACACGCTGAAGGTGACTGAGAA 45381  
QY 1500 GAACATAATAGATTCAATGAAAGCTTTGAGCCTTTATGGTAGAGCATGGCAGAGATTGAA 1559  
Db 45382 GAACATAATAGATTCAATGAAAGCTTTGAGCCTTTATGGTAGAGCATGGCAGAGATTGAA 45441  
QY 1560 GGTGATTTTATTTCCCTTTTATATCTCTTTATTTTTTGTGTTTGCAGAGATTGCTTCA 1619  
Db 45442 GGTGATTTTATTTCCCTTTTATATCTCTTTATTTTTTGTGTTTGCAGAGATTGCTTCA 45501

QY	1620	AACTGATTTGCTTTTTCATTTGGACAGAACATGTAGCAACAAAACTGCTGTCCAGAT	1679	QY	2700	ATGAAACTAAAAACCGTGGGGGAATAACAGGGAAGTCTCAGAAATAACAGGTAAGAAAGGAGT	2759
Db	45502	AACTGATTTGCTTTTTCATTTGGACAGAACATGTAGCAACAAAACTGCTGTCCAGAT	45561	Db	46581	ATGAAACTAAAAACCGTGGGGGAATAACAGGGAAGTCTCAGAAATAACAGGTAAGAAAGGAGT	46640
QY	1680	AAGAAGTCACGCTCAGAAATTTTCTCCAAAGTAAATCGGTTAAATTTTGAATGATGTT	1739	QY	2760	CAAACTCAGATTTGAATGCAAAATCTCTGAAAAACGGTAATGAGCAAGNACCTCAGACTT	2819
Db	45562	AAGAAGTCACGCTCAGAAATTTTCTCCAAAGTAAATCGGTTAAATTTTGAATGATGTT	45621	Db	46641	CAAACTCAGATTTGAATGCAAAATCTCTGAAAAACGGTAATGAGCAAGNACCTCAGACTT	46700
QY	1740	CTCATCTTCATTTGGCTTAATGCTTAAGACTTAATGAAAGCCAGGCAAGTTTTCGCTTCT	1799	QY	2820	ATCCGATGATATCCCTGCTAGTGGGAGCTCAATAACAAAGTTCTCTATCAC	2879
Db	45622	CTCATCTTCATTTGGCTTAATGCTTAAGACTTAATGAAAGCCAGGCAAGTTTTCGCTTCT	45681	Db	46701	ATCCGATGATATCCCTGCTAGTGGGAGCTCAATAACAAAGTTCTCTATCAC	46760
QY	1800	TTTGCTTTCTAGTCAGGAGATAGATAGATTAAGTTTATAGAGTTTATAGCAATAA	1859	QY	2880	ATCTCTCTTCAGAGCCAGATAGTATCCCAACACAGTTGCAGGAGATTAATCAGTCGTTTC	2939
Db	45682	TTTGCTTTCTAGTCAGGAGATAGATAGATTAAGTTTATAGAGTTTATAGCAATAA	45741	Db	46761	ATCTCTCTTCAGAGCCAGATAGTATCCCAACACAGTTGCAGGAGATTAATCAGTCGTTTC	46820
QY	1860	GTCTTAAAAATAGTTGAGAGAAATGACAGATGTAATCGTTTCTTTGTTTATGCCCTATAT	1919	QY	2940	CTAATCATATAATGTCAACCCCTTTTACAAACACCCGCTCTTTATCTGCCGCAACTTTCG	2999
Db	45742	GTCTTAAAAATAGTTGAGAGAAATGACAGATGTAATCGTTTCTTTGTTTATGCCCTATAT	45801	Db	46821	CTAATCATATAATGTCAACCCCTTTTACAAACACCCGCTCTTTATCTGCCGCAACTTTCG	46880
QY	1920	CTTGTTAAATCCAAACATGTATACATAGATTTCTTCAGAGAAATGTTAGTTTCTTTAGATTTC	1979	QY	3000	CCTCATCATTTTGGGCTCCCGATTTCTAGTGGTGGCTCACTGTTCCAGGGAACCTCACCTC	3059
Db	45802	CTTGTTAAATCCAAACATGTATACATAGATTTCTTCAGAGAAATGTTAGTTTCTTTAGATTTC	45861	Db	46881	CCTCATCATTTTGGGCTCCCGATTTCTAGTGGTGGCTCACTGTTCCAGGGAACCTCACCTC	46940
QY	1980	TTTCAGATAAATCTGTCTTCTTACCGATCTGAGGTAGTGGCAAAAGTGGGCTGAGTGC	2039	QY	3060	CGAATCTGGCTCCATGCGCCGAGCCACTGTTTGCAGCTGCTAGTGTGTTGGTGGGCTGCCA	3119
Db	45862	TTTCAGATAAATCTGTCTTCTTACCGATCTGAGGTAGTGGCAAAAGTGGGCTGAGTGC	45921	Db	46941	CGAATCTGGCTCCATGCGCCGAGCCACTGTTTGCAGCTGCTAGTGTGTTGGTGGGCTGCCA	47000
QY	2040	TAGAAATTTTGAATCTTCTTGTGATAAGCCATAGAGTTAAACCAATTTTTCATTTTCCA	2099	QY	3120	ATGGATTTATACCTTTATGTGCTCTCTTATGTTTCAAGTGGTTCACCTAGTCTCTCCCAT	3179
Db	45922	TAGAAATTTTGAATCTTCTTGTGATAAGCCATAGAGTTAAACCAATTTTTCATTTTCCA	45981	Db	47001	ATGGATTTATACCTTTATGTGCTCTCTTATGTTTCAAGTGGTTCACCTAGTCTCTCCCAT	47060
QY	2100	GTTCTGTCTATTTAACTCTGTTAGGTGCTATTAGATTTTGTGTTTACGTTTGTGTTAGA	2159	QY	3180	CTACTTTTGGACCATCATGTGATGTAGTGTACACAAAACAGCACTTTTACAACTATGGTT	3239
Db	45982	GTTCTGTCTATTTAACTCTGTTAGGTGCTATTAGATTTTGTGTTTACGTTTGTGTTAGA	46040	Db	47061	CTACTTTTGGACCATCATGTGATGTAGTGTACACAAAACAGCACTTTTACAACTATGGTT	47120
QY	2160	GGGTAAACAACTACTCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGA	2219	QY	3240	CTGTGACAGCCGAGAGCAAGAACACTCCGAGGCACTCAAAGGCTCGATCTTCACTGGACT	3299
Db	46041	GGGTAAACAACTACTCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGA	46100	Db	47121	CTGTGACAGCCGAGAGCAAGAACACTCCGAGGCACTCAAAGGCTCGATCTTCACTGGACT	47180
QY	2220	GCTATGGGTCAAGCGCTAGACATAGCTATTCTCTCCACGCTAAGCGTAAACCAAC	2279	QY	3300	CAGAGATGTTTCAAAATAAGAGTAAACCAAGTTTGTATCAGCAGCCCTTCTGCAACACTG	3359
Db	46101	GCTATGGGTCAAGCGCTAGACATAGCTATTCTCTCCACGCTAAGCGTAAACCAAC	46160	Db	47181	CAGAGATGTTTCAAAATAAGAGTAAACCAAGTTTGTATCAGCAGCCCTTCTGCAACACTG	47240
QY	2280	AATCTCTTATCTCGAAGACGGGAAGTGAACGATCTTTATGTCAAAAAACGGGTGTGAAT	2339	QY	3360	AGAGTGTATCAAAAGGTTTCAGATGGAGCAGAGACAGAAAAACAAAGTTGACCGCTCTCGT	3419
Db	46161	AATCTCTTATCTCGAAGACGGGAAGTGAACGATCTTTATGTCAAAAAACGGGTGTGAAT	46220	Db	47241	AGAGTGTATCAAAAGGTTTCAGATGGAGCAGAGACAGAAAAACAAAGTTGACCGCTCTCGT	47300
QY	2340	GATGAAAAGAGTCCCTTGATCAGAAAAAGTGTGCGATCTCTGAGGTGATTTTCATGGTC	2399	QY	3420	GTGGCTCAAAACACTCCGTGAGTGTGATGTTTGAAGCGGATGATCAGAAAAAGCAAG	3479
Db	46221	GATGAAAAGAGTCCCTTGATCAGAAAAAGTGTGCGATCTCTGAGGTGATTTTCATGGTC	46280	Db	47301	GTGGCTCAAAACACTCCGTGAGTGTGATGTTTGAAGCGGATGATCAGAAAAAGCAAG	47360
QY	2400	ATATGGCATCTTTTTCAGTGTGTACATTTGCTCTCTCATGTTTATTAATACAGATTGTGTG	2459	QY	3480	AGGATGGCAACCAATGTTGAGGTGAAAGAACCAATGAAGACACTTAATAACCTCAAACTT	3539
Db	46281	ATATGGCATCTTTTTCAGTGTGTACATTTGCTCTCTCATGTTTATTAATACAGATTGTGTG	46340	Db	47361	AGGATGGCAACCAATGTTGAGGTGAAAGAACCAATGAAGACACTTAATAACCTCAAACTT	47420
QY	2460	CTTCGTTTATAGATGCCAATGAAGATCGAACCAATCAAAAGCCTGAAGAGAAAACTCTG	2519	QY	3540	CAGAGTCCAAATCGACCCGCGCAGTAGAATCAGCTCCAAATATAACCGATCCTGAAAGCTG	3599
Db	46341	CTTCGTTTATAGATGCCAATGAAGATCGAACCAATCAAAAGCCTGAAGAGAAAACTCTG	46400	Db	47421	CAGAGTCCAAATCGACCCGCGCAGTAGAATCAGCTCCAAATATAACCGATCCTGAAAGCTG	47480
QY	2520	CAGGAAGCAACTGTTTCAGATTGTTTCACTCATCAGTATCTCTGCTGCTCCTCCATG	2579	QY	3600	TGCTCGACGAGTACTTACTTGGACTAAAGATCAACTTCTTTTATTTTCAAACTATTTTCT	3659
Db	46401	CAGGAAGCAACTGTTTCAGATTGTTTCACTCATCAGTATCTCTGCTGCTCCTCCATG	46460	Db	47481	TGCTCGACGAGTACTTACTTGGACTAAAGATCAACTTCTTTTATTTTCAAACTATTTTCT	47540
QY	2580	AATAAAGTTGTATAGACATCAAAACGCAAGCACTTTCCGCGAGTTCTTTCCTTTCACGG	2639	QY	3660	CATATAAATATTTGATCATTC-GGGTGCAATGTCCTTCCAGCTCTCTTCCAGAGAGT	3718
Db	46461	AATAAAGTTGTATAGACATCAAAACGCAAGCACTTTCCGCGAGTTCTTTCCTTTCACGG	46520	Db	47541	CATATAAATATTTGATCATTCGAAATTCGCTTCCAGCTCTCTTCCAGAGAGT	47600
QY	2640	GAAGAGTAAAAAAAATCTTTTCAATGCTATTGAGGTTTAAAGACGATTTAGTACTTTTC	2699	QY	3719	ATTGCGCAAAAGTTTTTACATATCGAGAGAAACACAGAGAGGAGAACACAAACAGA	3778
Db	46521	GAAGAGTAAAAAAAATCTTTTCAATGCTATTGAGGTTTAAAGACGATTTAGTACTTTTC	46580	Db	47601	ATTGCGCAAAAGTTTTTACATATCGAGAGAAACACAGAGAGGAGAACACAAACAGA	47660
				QY	3779	ACAAAGATATCCAATGGCACTTTGATCTTTAACTTTCAGCTCAGTTTAAACCAAGGTTGATGA	3838





Db	121		123
Qy	1255	ACTTGATGTTTTTGTGTTTGTGGATTAGTGGAAATTTGTAAAGAGAAATCTGAAGTTG	1314
Db	124	-----AGTGGAAATTTGTAAAGAGNAGATCTGAAGTTG	156
Qy	1315	TGTAGAGGAGCTTAGTGATGGAGACAAATTCGTCTGGAGAAAGATCTGGTTATTAAAGGTAA	1374
Db	157	TGTAGAGGAGCTTAGTGATGGAGACAAATTCGTCTGGAGAAAGATCTGGTTATTAA	210
Qy	1375	ATTAACTAAATTTTAGGGGAAGATGATTGTTTAGTGTCAAAGATTGAGAATTTTAAAT	1434
Db	211	-----	210
Qy	1435	GAAACTTGATATAGACTCGGAAGCCATATACGATAACAAAGCAACGTAAGAGGTGGACTG	1494
Db	211	-----AGACTCGGAAGCCATATACGATAACAAAGCAACGTAAGAGGTGGACTG	258
Qy	1495	AGGAAGCAATAATAGATTCAATTTGAAGCTTTGAGGCTTTATGGTAGAGCATGGCAGAAGA	1554
Db	259	AGGAAGCAATAATAGATTCAATTTGAAGCTTTGAGGCTTTATGGTAGAGCATGGCAGAAGA	318
Qy	1555	TTGAAGGTTGATTTTATTTCCCTTTATATGTCCTTATTTTGTGTTGTCAGAGGTTTGT	1614
Db	319	TTGA	322
Qy	1615	CTTCAAACTGATTTTGCTTTTTTTCATTTGGACAGAACATGTAGCAACAAAAACTGCTGTC	1674
Db	323	-----AGACATGTAGCAACAAAAACTGCTGTC	350
Qy	1675	CAGATAAGAAAGTCACGCTCAGAAATTTTCTCAAGGTAAATTCGGTTAAATTTGAAATG	1734
Db	351	CAGATAAGAAAGTCACGCTCAGAAATTTTCTCAAGGTAAATTCGGTTAAATTTGAAATG	410
Qy	1735	ATGTTCTCATCTTCATTCGTTAAATGCTTAAGACTTTATTCAGAGCCAGGCAAGTTTCTG	1794
Db	411	ATGTTCTCATCTTCATTCGTTAAATGCTTAAGACTTTATTCAGAGCCAGGCAAGTTTCTG	470
Qy	1795	CTTCTTTTGTCTTCTTAGTCAGAGAGATAGATAGATTACGTTTTTATAGATTTTAGTAATGAGC	1854
Db	471	CTTCTTTTGTCTTCTTAGTCAGAGAGATAGATAGATTACGTTTTTATAGATTTTAGTAATGAGC	530
Qy	1855	AATAAGTCTTAAATAGTTGGAGAAATGACGAGATGTAATCGTTTCTTTTGTGTTATGTC	1914
Db	531	AATAAGTCTTAAATAGTTGGAGAAATGACGAGATGTAATCGTTTCTTTTGTGTTATGTC	590
Qy	1915	TATATCTTGTGTTAATCCACAAACATGTACATAGATTCTTCAGAGAAATGTTAGTTTCTTTA	1974
Db	591	TATATCTTGTGTTAATCCACAAACATGTACATAGATTCTTCAGAGAAATGTTAGTTTCTTTA	650
Qy	1975	GATTTCTCAGATAAACTTGTGCTCTCTTACCGATTCTGAGGTAGTGGCAAAAGTGGGCTG	2034
Db	651	GATTTCTCAGATAAACTTGTGCTCTCTTACCGATTCTGAGGTAGTGGCAAAAGTGGGCTG	710
Qy	2035	AGTGTAGAAAATTTTGAATGTTCTTGATTAAGCCATAGAGGTAAACCAATTTTGAAT	2094
Db	711	AGTGTAGAAAATTTTGAATGTTCTTGATTAAGCCATAGAGGTAAACCAATTTTGAAT	770
Qy	2095	TTCAGGTTCTGTCATTTAACTTGTGTTAGGTGTCATTAGATTGTTTGTGTTTACGTTTGT	2154
Db	771	TTCAGGTTCTGTCATTTAACTTGTGTTAGGTGTCATTAGATTGTTTGTGTTTACGTTTGT	829
Qy	2155	TTAGAGGGTAACAAAATCTACTCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTTAAAG	2214
Db	830	TTAGAGGGTAACAAAATCTACTCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTTAAAG	889
Qy	2215	GTGTAGCTATGGTCAAGGCTAGACATAGCTTATCTCTCCACGGCTTAAGCGTTAAAC	2274
Db	890	GTGTAGCTATGGTCAAGGCTAGACATAGCTTATCTCTCCACGGCTTAAGCGTTAAAC	949
Qy	2275	CAAAACAATCTTATCTCGAAAGACGGGAAGTGGAAACGATCCTTATGTCAAAACCGGTC	2334
Db			
Db	950	CAAAACAATCTTATCTCGAAAGACGGGAAGTGGAAACGATCCTTATGTCAAAACCGGTC	1009
Qy	2335	TGAATGATGAAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTCGAGGTGATTTTCA	2394
Db	1010	TGAATGATGAAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTCGATCCTG	1057
Qy	2395	TGTCATATGGCATCTTTTTCAGTGTGTGCATTTGCTCTCTCATGTTATTAAATACAGATT	2454
Db	1058	-----	1057
Qy	2455	GTGTGCTTCGTTTATAGATGGCCCAATCAAGATCGACAACAATCAAGCCCTGAAGAGAAA	2514
Db	1058	-----AGATGGCCCAATCAAGATCGACAACAATCAAGCCCTGAAGAGAAA	1102
Qy	2515	CTCTGAGGAAGACAACTGTTTCAGATTGTTTCTACTCATCATGATCTCTCTGCTGCATCCT	2574
Db	1103	CTCTGAGGAAGACAACTGTTTCAGATTGTTTCTACTCATCATGATCTCTCTGCTGCATCCT	1162
Qy	2575	CCATGAATAAAGTTGTATAGACATCAAAACCAAGCACCTTCCGCGAGTTCTTGCTT	2634
Db	1163	CCATGAATAAAGTTGTATAGACATCAAAACCAAGCACCTTCCGCGAGTTCTTGCTT	1222
Qy	2635	CACGGGAAGAGGTAAAAAACAATCTTTCATTGCTATTGAGGTTTTAAGACGATTAGTAC	2694
Db	1223	CACGGGAAG	1231
Qy	2695	TTTTTCATGAATAAAACCGTGGGGAATAACAGGAAGTCAGAATAACAGGGTAAGAAA	2754
Db	1232	-----AGGGAAGTCAGNAATAACAGGGTAAGAAA	1259
Qy	2755	GGAGTCAAACTCAGATTTTGAATGCAAAATCTCTGGAAAACGGTAATGAGCAAGGACCTCA	2814
Db	1260	GGAGTCAAACTCAGATTTTGAATGCAAAATCTCTGGAAAACGGTAATGAGCAAGGACCTCA	1319
Qy	2815	GACTTATCCGATGCATATCCCTGTGTAGTGCATTTGGGGAGCTCAATAACAAGTTCTCT	2874
Db	1320	GACTTATCCGATGCATATCCCTGTGTAGTGCATTTGGGGAGCTCAATAACAAGTTCTCT	1379
Qy	2875	ATCACATCTCTCTTCAGAGCCAGATAGTCATCCCCACACAGTTGCGAGAGATTATCAGTC	2934
Db	1380	ATCACATCTCTCTTCAGAGCCAGATAGTCATCCCCACACAGTTGCGAGAGATTATCAGTC	1439
Qy	2935	GTTTCTCTAATCATATATATGTCAAACCCCTTTTACAAACACCGGCTCTTTATATCTGCGCAAC	2994
Db	1440	GTTTCTCTAATCATATATATGTCAAACCCCTTTTACAAACACCGGCTCTTTATATCTGCGCAAC	1499
Qy	2995	TTTTCGGCTCATCATTTTGGCTCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTC	3054
Db	1500	TTTTCGGCTCATCATTTTGGCTCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTC	1559
Qy	3055	ACCTCCGAATCTGGCTGCCATGGCCGAGCCACTGTTGCAAGCTGCTAGTGTCTGGTGGGC	3114
Db	1560	ACCTCCGAATCTGGCTGCCATGGCCGAGCCACTGTTGCAAGCTGCTAGTGTCTGGTGGGC	1619
Qy	3115	TGCAATGGATTATTAACCTTTTATGTCCTCTTATGTTTCAAGTGGTGGTTCCTAGTGTCTC	3174
Db	1620	TGCAATGGATTATTAACCTTTTATGTCCTCTTATGTTTCAAGTGGTGGTTCCTAGTGTCTC	1679
Qy	3175	TCCATCTACTTTTGGACCATCATGTCATGTAGAGTACACAAAAGCAAGCACTTTTACAACA	3234
Db	1680	TCCATCTACTTTTGGACCATCATGTCATGTAGAGTACACAAAAGCAAGCACTTTTACAACA	1739
Qy	3235	TGGTTCCTGTGACAGCCGAGAGCAAGCACTCCGAGGCACTCAAGGCTCGATCTTCACT	3294
Db	1740	TGGTTCCTGTGACAGCCGAGAGCAAGCACTTCGAGGCACTCAAGGCTCGATCTTCACT	1799
Qy	3295	GGACTCAGAGGATGTTGAAAAAAGAGTAAACAGTTTGTTCATGAGCAGCTTCTGCAAC	3354
Db	1800	GGACTCAGAGGATGTTGAAAAAAGAGTAAACAGTTTGTTCATGAGCAGCTTCTGCAAC	1859
Qy	3355	ACCTGAGATGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAAACAAGTTGACCGGTC	3414
Db	1860	ACCTGAGATGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAAACAAGTTGACCGGTC	1919





QY 2840 CTAGTGCATTTGGGAGCTCAATTAACAAGTTTCTCTATCAATCCTCTCTCAGAGCCAGAT 2899  
Db CTAGTGCATTTGGGAGCTCAATTAACAAGTTTCTCTATCAATCCTCTCTCAGAGCCAGAT 990  
QY 2900 AGTCATCCACACAGTTGAGGAGATTATCAGTCGTTTCTCTATCAATCCTCTCTCAGAGCCAGAT 2959  
Db AGTCATCCACACAGTTGAGGAGATTATCAGTCGTTTCTCTATCAATCCTCTCTCAGAGCCAGAT 1050  
QY 2960 CTTTACAACACCGGCTCTTTATATCTGCGCAACTTTGCGCTCATCATTTTGGCCCTCCC 3019  
Db CTTTACAACACCGGCTCTTTATATCTGCGCAACTTTGCGCTCATCATTTTGGCCCTCCC 1110  
QY 3020 GATCTAGTGGTGGCTCAGCTGTTTCCAGGGAACCTCACCTCGAATCTGGCTGCGCATGGCC 3079  
Db GATCTAGTGGTGGCTCAGCTGTTTCCAGGGAACCTCACCTCGAATCTGGCTGCGCATGGCC 1170  
QY 3080 GCAGCCACTGTTGAGCTGCTAGTGTCTTGGTGGCTGCGCAATGATTTACCTTTATGT 3139  
Db GCAGCCACTGTTGAGCTGCTAGTGTCTTGGTGGCTGCGCAATGATTTACCTTTATGT 1230  
QY 3140 GCTCCTCTTAGTTTCAAGTGGTTTCACTAGTATCATCTCTCATCTACTTTTGGACCATCATGT 3199  
Db GCTCCTCTTAGTTTCAAGTGGTTTCACTAGTATCATCTCTCATCTACTTTTGGACCATCATGT 1290  
QY 3200 GATGTAGATACACAAAAGCAAGCACTTTACAACATGGTTCTGTGAGAGCCGAGAGCAA 3259  
Db GATGTAGATACACAAAAGCAAGCACTTTACAACATGGTTCTGTGAGAGCCGAGAGCAA 1350  
QY 3260 GAACACTCCGAGGATCAAAAGGCTTCATCTTCACTGGACTCAGAGGATGTTGAAATTAAG 3319  
Db GAACACTCCGAGGATCAAAAGGCTTCATCTTCACTGGACTCAGAGGATGTTGAAATTAAG 1410  
QY 3320 AGTAAACAGTTTGTCTATGAGCGCTTCTGCAACACTGAGAGTGATGCAAGGGTTCA 3379  
Db AGTAAACAGTTTGTCTATGAGCGCTTCTGCAACACTGAGAGTGATGCAAGGGTTCA 1470  
QY 3380 GATGGAGCAGGACAGAGAAACAAAGTTGACGGTCTCTGCTGCTCAAACTCCGCTCG 3439  
Db GATGGAGCAGGACAGAGAAACAAAGTTGACGGTCTCTGCTGCTCAAACTCCGCTCG 1530  
QY 3440 AGTAGTATCATGTTGAGCGGATGATCATCAAGAGCAAGAGATGCAACCAATGGTGAG 3499  
Db AGTAGTATCATGTTGAGCGGATGATCATCAAGAGCAAGAGATGCAACCAATGGTGAG 1590  
QY 3500 GTGAAAGAAACGAATGAAGACACTTAATAACCTCAAACTTCAAGTCCAAATGCACGCCGC 3559  
Db GTGAAAGAAACGAATGAAGACACTTAATAACCTCAAACTTCAAGTCCAAATGCACGCCGC 1650  
QY 3560 AGTAGAATCAGCTCCAAATATAACCGATCCATGGAAGTCTGTGCTGACGAGGTACTTACT 3619  
Db AGTAGAATCAGCTCCAAATATAACCGATCCATGGAAGTCTGTGCTGACGAGGTACTTACT 1700  
QY 3620 TGGACTAAAGATCAACTCTCTTTTATTTCAAACTCATTTTCTCATATAAATATTGTACATTC 3679  
Db TGGACTAAAGATCAACTCTCTCTTTTATTTCAAACTCATTTTCTCATATAAATATTGTACATTC 1701  
QY 3680 GGGTCGAATTCCTTCCAAGCTCTCTTCTCCAGAGAGGATTGCGCGAAAGTTTACATA 3739  
Db GGGTCGAATTCCTTCCAAGCTCTCTTCTCCAGAGAGGATTGCGCGAAAGTTTACATA 1760  
QY 3740 TCAGAGAAACACAGAGAGAGAGAAACAAACAAACAAAGAAACAAAGATATCCAAATGGCACT 3799  
Db TCAGAGAAACACAGAGAGAGAGAGAAACAAACAAACAAAGAAACAAAGATATCCAAATGGCACT 1820  
QY 3800 TGATCTTAACCTTCAACAGCTCAGTTTAAACACAGTTGATGATCAAGAGAGAGAGAAACAC 3859  
Db TGATCTTAACCTTCAACAGCTCAGTTTAAACACAGTTGATGATCAAGAGAGAGAGAAACAC 1880  
QY 3860 AGGATTTCTTGGAAATCGGATTAGATGCTTCAAAAGCTTAATGAGTAGAGGAGAAACAGGTTT 3919  
Db AGGATTTCTTGGAAATCGGATTAGATGCTTCAAAAGCTTAATGAGTAGAGGAGAAACAGGTTT 1940

QY 3920 TAAACCATACAAAAGATGTTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAAATCC 3979  
Db TAAACCATACAAAAGATGTTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAAATCC 2000  
QY 3980 TATCATTTATGTCGAAACAGAAAGATCCCAACCGATCGGTTGAAACTCAAGCTTCCAC 4039  
Db TATCATTTATGTCGAAACAGAAAGATCCCAACCGATCGGTTGAAACTCAAGCTTCCAC 2060  
QY 4040 ATGAGACTTATTTTCACTCTGATCTGTTGTTTGTACTCTGTTTAAAGTTTCAAGACCA 4099  
Db ATGAGACTTATTTTCACTCTGATCTGTTGTTTGTACTCTGTTTAAAGTTTCAAGACCA 2120  
QY 4100 CTGCTACATTTTCTTTTCTTTTGGAGCCCTTTGATTTGTTTCTTGTCCATAGTCTTCC 4159  
Db CTGCTACATTTTCTTTTCTTTTGGAGCCCTTTGATTTGTTTCTTGTCCATAGTCTTCC 2180  
QY 4160 TGTAAACATTTGACTCTGATTTATTTCAACAAATCATAACTGTTTAACTTTTTTTTCCA 4219  
Db TGTAAACATTTGACTCTGATTTATTTCAACAAATCATAACTGTTTAACTTTTTTTTCCA 2240  
QY 4220 ACCTGGAAAGAA 4231  
Db AAAAAAAAAAAAAA 2252

RESULT 6  
ATU28422  
LOCUS  
DEFINITION  
Arabidopsis thaliana DNA-binding protein CCA1 (CCA1) mRNA, complete cds.  
ACCESSION  
U28422  
VERSION  
U28422.1 GI:1777442  
KEYWORDS  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 (bases 1 to 2240)  
Wang,Z.Y., Kenigsbuch,D., Sun,L., Harel,E., Ong,M.S. and Tobin,E.M.  
A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene  
Plant Cell 9 (4), 491-507 (1997)  
97290202  
MEDLINE  
PUBMED  
9144958  
REFERENCE  
2 (bases 1 to 2240)  
Wang,Z.  
Direct Submission  
Submitted (06-JUN-1995) Zhiyong Wang, Dept. of Biology, University of California at Los Angeles, 405 Hillgard Ave., Los Angeles, CA 90095, USA  
JOURNAL  
TITLE  
Arabidopsis thaliana DNA-binding protein CCA1 (CCA1) mRNA, complete cds.  
AUTHORS  
Wang,Z.  
JOURNAL  
TITLE  
Arabidopsis thaliana DNA-binding protein CCA1 (CCA1) mRNA, complete cds.  
FEATURES  
source  
1. .2240  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/dev\_stage="21-day-old green seedlings"  
1. .2240  
/genes="CCA1"  
238. .2064  
/genes="CCA1"  
/function="DNA binding protein"  
/note="myb-related DNA binding protein"  
/codon\_start=1  
/product="CCA1"  
/protein\_id="AAB40525.1"  
/db\_xref="GI:1777443"  
/translation="METNSGDELVIKTRPYITKQERWTEENRFRLEALRLVGR  
ANQKEEHVATKTVQIRSHQKFFSKVEAEAKGVAMQALDIAPPRPKRPN  
PYRKTGSGTILMSKTGVNDKESLGSEKVSHPEDMANEDRQSKPEKTLQEDKSDC  
FTHYLSAASSKCIETSNASTREFLPRESQNNRVKESNSDLNAKSLNGN  
EQGPQTPMHIPLVPLGSSITSSLSHPSPDPDHPHTVAGDQVSPFNHIMSTLLQTP

```
ALYTAATFASFMPPDSSGGSPVPGNSPPNLAAMAAATVAAASAMWAANGLLPLCAPL
SSGGFTSHPPSTTPGSCDVEYTKASTLQHGVSQSEAEKASRSLDSDEVENKS
KPYCHEOPSATPSSDAKSGDAGDKOVDRSSCGSNTPSSSDVDEADASERQEDTNG
EVKETNEDTNKPTDSKSNARRSRISNITDPMKSVSDGRIAFQALFSRVLQSFY
REHREBEQQQQRYPMALDLNFTLAQLTVPDQBEKRNTGFLIGLADSKLMSRGT
GFAPYKCSMEAKESRLNNPIIHVEQKPKRMURJETQAST"
2240
polyA_site
/gene="CCAl"
/note="14 A nucleotides"
BASE COUNT 685 a 473 c 477 g 605 t
ORIGIN
Query Match 35.1%; Score 1525.2; DB 8; Length 2240;
Best Local Similarity 88.1%; Pred. No. 5.7e-306;
Matches 1797; Conservative 0; Mismatches 3; Indels 240; Gaps 3;
QY 2180 TCTCTCTCAGGTAGAGAAAGGCTGAAGCTAAAGGTGTAAGTCTAGCTTGGTCAAGCGCTAGA 2239
DB 441 TTTCTCAAGGTAGAGAAAGGCTGAAGCTAAAGGTGTAAGTCTAGCTTGGTCAAGCGCTAGA 500
QY 2240 CATAGCTATTCTCTCCACGGCTAAAGCGTAAACCAACAATCTTATCTCGAAAGAC 2299
DB 501 CATAGCTATTCTCTCCACGGCTAAAGCGTAAACCAACAATCTTATCTCGAAAGAC 560
QY 2300 GGAAGTGGAAACGATCTTATGTCAAAAACGGGTGTAAGTATGATGGAAGAGTCCCTTGG 2359
DB 561 GGAAGTGGAAACGATCTTATGTCAAAAACGGGTGTAAGTATGATGGAAGAGTCCCTTGG 620
QY 2360 ATCAGAAAAAGTGTGCGATCTCAGGTGATTTTTCATGGTCATATGGCATCTTTTGCAGT 2419
DB 621 ATCAGAAAAAGTGTGCGATCTCCTG----- 643
QY 2420 GTGTCAATTTGCTCTCATGTTTATATACAGATTGTGTGCTTCGTTTATAGATGGCAA 2479
DB 644 -----AGATGGCAA 653
QY 2480 TGAAGATCGACAACAATCAAGCTGAGAGAAAACTCTCAGAGAGACAATCTGTTTACAGA 2539
DB 654 TGAAGATCGACAACAATCAAGCTGAGAGAAAACTCTCAGAGAGACAATCTGTTTACAGA 713
QY 2540 TTGTTTCACTCATCAGTATCTCTGCTGCATCTCTCATCAATAAAGTTGTATAGAGAC 2599
DB 714 TTGTTTCACTCATCAGTATCTCTGCTGCATCTCTCATCAATAAAGTTGTATAGAGAC 773
QY 2600 ATCAACGCAAGCACTTTCCGAGTCTTTCGCTTCAACGGGAAGAGGTAAAAACAATCT 2659
DB 774 ATCAACGCAAGCACTTTCCGAGTCTTTCGCTTCAACGGGAAG----- 817
QY 2660 TTCATTGCTATTGAGGTTTAAAGACGATTAGTACTTTTCATGAAACTAAACCGTGGG 2719
DB 818 ----- 817
QY 2720 GAATAACAGGGAAGTCAGATAACAGGTAAGAAAGAGTCAAACTCAGATTTGAATGCA 2779
DB 818 -----AGGGAAGTCAGATAACAGGTAAGAAAGAGTCAAACTCAGATTTGAATGCA 870
QY 2780 AAATCTCTGMAAACGGTAATGAGCAAGACCTCAGACTTATCCGATGATATCCCTGTG 2839
DB 871 AAATCTCTGMAAACGGTAATGAGCAAGACCTCAGACTTATCCGATGATATCCCTGTG 930
QY 2840 CTAGTGCCATTGGGAGCTCAATAACAAGTTCTCTATACATCTCTCTCAGAGCCAGAT 2899
DB 931 CTAGTGCCATTGGGAGCTCAATAACAAGTTCTCTATACATCTCTCTCAGAGCCAGAT 990
QY 2900 AGTCATCCCCACACAGTTGAGGAGATTATCAGTCGTTTCTTAACTATATATATGTCACAC 2959
DB 991 AGTCATCCCCACACAGTTGAGGAGATTATCAGTCGTTTCTTAACTATATATGTCACAC 1050
QY 2960 CTTTTCACAAACCGGCTCTTATATCTGCGGAACTTTTCCGCTCATCATTTTGGCCTCCC 3019
DB 1051 CTTTTCACAAACCGGCTCTTATATCTGCGGAACTTTTCCGCTCATCATTTTGGCCTCCC 1110
QY 3020 GATTCTAGTGGTGGCTCACTGTTCAGGGAACCTCACCTCCGAACTCTGCGTGGCTGGCC 3079
```

```
||||| GATTCTAGTGGTGGCTCACCTGTTCCAGGAACTCACCTCGAATCTGGCTGCCATGCC 1170
||||| GCAGCCACTGTTGCAGCTGCTAGTGTGTTGGTGGCTGCCAATGGAATATTACCTTTATGT 3139
||||| GCAGCCACTGTTGCAGCTGCTAGTGTGTTGGTGGCTGCCAATGGAATATTACCTTTATGT 1230
||||| GCTCTCTTAGTTTACAGGTGGTTTCACTAGTATCATCTCCATCTACTTTTGGACCATCATGT 3199
||||| GCTCTCTTAGTTTCAAGGTGGTTTCACTAGTATCATCTCCATCTACTTTTGGACCATCATGT 1290
||||| GATGTAGAGTACACAAAAGCAAGCACTTTTACAACATGGTTCTGTGCAGAGCCGAGAGCAA 3259
||||| GATGTAGAGTACACAAAAGCAAGCACTTTTACAACATGGTTCTGTGCAGAGCCGAGAGCAA 1350
||||| GAACACTCCAGGCAATCAAAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAAATAAG 3319
||||| GAACACTCCAGGCAATCAAAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAAATAAG 1410
||||| AGTAAACCAAGTTTGTATGAGCAGCTTCTGCAACACCTCTGAGAGTGATGCAAGGGTTCA 3379
||||| AGTAAACCAAGTTTGTATGAGCAGCTTCTGCAACACCTCTGAGAGTGATGCAAGGGTTCA 1470
||||| GATGGAGCAGGAGACAGAAAAAAGTTGACCGGTCTCTGTGGCTCAAAACACTCCGCTCG 3439
||||| GATGGAGCAGGAGACAGAAAAAAGTTGACCGGTCTCTGTGGCTCAAAACACTCCGCTCG 1530
||||| AGTAGTGTATGTTTGGGCGGATGCATCAGAAAGCGCAAGAGGATGCAATGGTGAG 3499
||||| AGTAGTGTATGTTTGGGCGGATGCATCAGAAAGCGCAAGAGGATGCAATGGTGAG 1590
||||| GTCAAAAGAAACGAATGAAGACACTTAATAAACCTCAAACTTCAGAGTCCCAATGCACGCCG 3559
||||| GTCAAAAGAAACGAATGAAGACACTTAATAAACCTCAAACTTCAGAGTCCCAATGCACGCCG 1650
||||| AGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGTCTGACGAGTACTTACT 3619
||||| AGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGTCTGACGAGTACTTACT 1700
||||| TGGACTTAAGATCAACTCTCTTTATTTCAAATCATTTTCTCATATAAATATTGTACATTC 3679
||||| ----- 1700
||||| GGTCGAATTGCTTCCAAAGCTCTTCTCCAGAGAGGTATTGCGCGCAAGTTTTCACATA 3739
||||| GGTCGAATTGCTTCCAAAGCTCTTCTCCAGAGAGGTATTGCGCGCAAGTTTTCACATA 1760
||||| TCAGAAAGAACACAGAGAGGAGAAACAAACAAACAAAGAAATATCCAATGGCACT 3799
||||| TCAGAAAGAACACAGAGAGGAGAAACAAACAAACAAAGAAATATCCAATGGCACT 1820
||||| TGATCTTAACCTTACAGCTCAGTTTAACACAGTTTGTATGATCAAGAGAGAGAGAAACAC 3859
||||| TGATCTTAACCTTACAGCTCAGTTTAACACAGTTTGTATGATCAAGAGAGAGAGAAACAC 1880
||||| AGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAAGCTTAATGAGTAGAGAGAGAAACAGGTTT 3919
||||| AGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAAGCTTAATGAGTAGAGAGAGAAACAGGTTT 1940
||||| TAAACCATACAAAGATGTTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACCAATCC 3979
||||| TAAACCATACAAAGATGTTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACCAATCC 2000
||||| TATCATTTGAGTGGACAGAAAGATCCCAACCGATGCGGTTGGAAACTCAAGCTTCCAC 4039
||||| TATCATTTGAGTGGACAGAAAGATCCCAACCGATGCGGTTGGAAACTCAAGCTTCCAC 2060
||||| ATGAGACTCTATTTTTCATCTGATCTGTTGTTGATCTGTTTAAAGTTTTCAGACCA 4099
||||| ATGAGACTCTATTTTTCATCTGATCTGTTGTTGATCTGTTTAAAGTTTTCAGACCA 2120
||||| CTGTACATTTTCTTTTCTTTGAGGCTTTGTATTTGTTTCTTTGCTCATAGTCTCC 4159
```

Db 2121 CTGCTACATTTTCTTTCTTTTGAGCCCTTTGTATTTGTTCTTGTCTCATAGTCTTCC 2180

Qy 4160 TGTAACATTGACTCTGTATTTATTCACAAATCATAAACTGTTTAATCTTTTTTTTCCA 4219

Db 2181 TGTAACATTGACTCTGTATTTATTCACAAATCATAAACTGTTTAATCTTTTTTTTCCA 2240

RESULT 7

AX506805

LOCUS 1827 bp DNA linear PAT 27-SEP-2002

DEFINITION Sequence 1500 from Patent WO0216655.

ACCESSION AX506805

VERSION AX506805.1 GI:233988042

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1

REFERENCE Harper, J.F., Kreps, J.J., Wang, X. and Zhu, T.

AUTHORS Stress-regulated genes of plants, transgenic plants containing

TITLE same, and methods of use

JOURNAL Patent: WO 0216655-A 1500 28-FEB-2002;

The Scripps Research Institute (US) ; Syngenta Participations AG (CH)

FEATURES

source Location/Qualifiers

1..1827

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/db\_xref="taxon:3702"

BASE COUNT 591 a 397 c 419 g 420 t

ORIGIN

Query Match 31.1%; Score 1349.2; DB 6; Length 1827;

Best Local Similarity 87.0%; Pred. No. 2e-269;

Matches 1621; Conservative 0; Mismatches 3; Indels 240; Gaps 3;

Qy 2180 TCTCTCTCAGGTAGAGAAAGAGCTGAAGCTAAAGGTGTAGCTATGGTCAAGCGCTAGA 2239

Db 204 TTTCTCAAAGGTAGAGAAAGAGCTGAAGCTAAAGGTGTAGCTATGGTCAAGCGCTAGA 263

Qy 2240 CATAGCTATTCCTCCACGGCTAAGCTTAACCAACAACTCTTATCCTCGAAGAC 2299

Db 264 CATAGCTATTCCTCCACGGCTAAGCTTAACCAACAACTCTTATCCTCGAAGAC 323

Qy 2300 GGAAGTGGAAACGATCCTTATGTCAAAAACGGGTGTGAATGATGGAAGAGTCCCTTGG 2359

Db 324 GGAAGTGGAAACGATCCTTATGTCAAAAACGGGTGTGAATGATGGAAGAGTCCCTTGG 383

Qy 2360 ATCAGAAAAGTGTGCGATCCTCGAGGTGATTTTCATGGTCAATATGGCATCTTTTTCAGT 2419

Db 384 ATCAGAAAAGTGTGCGATCCTG----- 406

Qy 2420 GTGTACATTTGCTCCTCATGTTATTAATACAGATTGTGTCTCGTTTATAGATGCCAA 2479

Db 407 -----AGATGGCCAA 416

Qy 2480 TGAAGATCGACAACAATCAAAGCTGAAGAGAAAACCTGCGAGAAAGACAACCTGTTTACA 2539

Db 417 TGAAGATCGACAACAATCAAAGCTGAAGAGAAAACCTGCGAGAAAGACAACCTGTTTACA 476

Qy 2540 TTGTTTCACTCATCATGATATCTCTCTGTCGATCCTCCATGAATPAAAAGTTGTATAGAGAC 2599

Db 477 TTGTTTCACTCATCATGATATCTCTCTGTCGATCCTCCATGAATPAAAAGTTGTATAGAGAC 536

Qy 2600 ATCAAACGAAGCAGCTTTCCGGAGTTCTTGCTTCCACGGAGAGGTAAAAACAATCT 2659

Db 537 ATCAAACGAAGCAGCTTTCCGGAGTTCTTGCTTCCACGGAGAGGTAAAAACAATCT 580

Qy 2660 TTCATTGCTATTGAGGTTTTAAGACGATTAGTACTTTTCATGAAACTAAACCGGTGGG 2719

Db 581 ----- 580

Qy 2720 GAATAACAGGGAAGTCAGAAATAACAGGGTAAGAAAGAGTCAAACTCAGATTTGAATGCA 2779

Db 581 -----AGGGAAGTCAGAAATAACAGGGTAAGAAAGAGTCAAACTCAGATTTGAATGCA 633

Qy 2780 AATCTCTGGAANAACGGTAATGAGCAAGGACCTCAGACTTATCGGATGCATATCCCTGTG 2839

Db 634 AATCTCTGGAANAACGGTAATGAGCAAGGACCTCAGACTTATCGGATGCATATCCCTGTG 693

Qy 2840 CTAGTGCCATTGGGAGCTCAATAACAAAGTTCTCTATCACATCTCTCTTCAGAGCCAGAT 2899

Db 694 CTAGTGCCATTGGGAGCTCAATAACAAAGTTCTCTATCACATCTCTCTTCAGAGCCAGAT 753

Qy 2900 AGTCATCCCCACACAGTTCGAGGAGTATATCAGTCGTTTCTTAATCATATAATGTCACCC 2959

Db 754 AGTCATCCCCACACAGTTCGAGGAGTATATCAGTCGTTTCTTAATCATATAATGTCACCC 813

Qy 2960 CTTTTCACAAACACCGGCTCTTTTACTGCGGAACTTTTCGCCCTCATCATTTTGGCCTCCC 3019

Db 814 CTTTTCACAAACACCGGCTCTTTTACTGCGGAACTTTTCGCCCTCATCATTTTGGCCTCCC 873

Qy 3020 GATTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACCTCCGAATCTGGCTGCCATGCC 3079

Db 874 GATTCTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACCTCCGAATCTGGCTGCCATGCC 933

Qy 3080 GCAGCCACTGTTGCAGCTGCTAGTGTGGTGGGCTGCCAATGGATTTATACCTTTATGT 3139

Db 934 GCAGCCACTGTTGCAGCTGCTAGTGTGGTGGGCTGCCAATGGATTTATACCTTTATGT 993

Qy 3140 GCTCCTCTTAGTTAGTGGTGGTTCACCTAGTATCTCTCCATCTACTCTTTTGGACCATCATGT 3199

Db 994 GCTCCTCTTAGTTAGTGGTGGTTCACCTAGTATCTCTCCATCTACTCTTTTGGACCATCATGT 1053

Qy 3200 GATGTAGGTACACAAAAGCAAGCACTTTTACACATGGTTCCTGTGCAGAGCCGAGAGCAA 3259

Db 1054 GATGTAGGTACACAAAAGCAAGCACTTTTACACATGGTTCCTGTGCAGAGCCGAGAGCAA 1113

Qy 3260 GAACACTCCGAGGATCAAAAGGCTCGATCTTCACTGGACTCAGAGAGATGTTGAAAATAAG 3319

Db 1114 GAACACTCCGAGGATCAAAAGGCTCGATCTTCACTGGACTCAGAGAGATGTTGAAAATAAG 1173

Qy 3320 AGTAAACCAAGTTTGTATGAGCAGCCTTCTGCAACACCTGAGAGTGTGATGAGGTTCA 3379

Db 1174 AGTAAACCAAGTTTGTATGAGCAGCCTTCTGCAACACCTGAGAGTGTGATGAGGTTCA 1233

Qy 3380 GATGGAGCAGGAGACAGAAAACAAAGTTGACCGGCTCTCGTGTGCTCAACACCTCCGTCG 3439

Db 1234 GATGGAGCAGGAGACAGAAAACAAAGTTGACCGGCTCTCGTGTGCTCAACACCTCCGTCG 1293

Qy 3440 AGTAGTGATGATGTTGAGGCGGATGCATCAGAAAGCAAGAGGATGGCACTAATGTTGAG 3499

Db 1294 AGTAGTGATGATGTTGAGGCGGATGCATCAGAAAGCAAGAGGATGGCACTAATGTTGAG 1353

Qy 3500 GTGAAAGAAACGAATGAAGACACTAATAAACCCTCAAACTTTCAGAGTCCAATGACGCGCG 3559

Db 1354 GTGAAAGAAACGAATGAAGACACTAATAAACCCTCAAACTTTCAGAGTCCAATGACGCGCG 1413

Qy 3560 AGTAGAATCAGCTCCCAATATACCGATCCATGAACTCTGTGTCGACGAGGTACTTACT 3619

Db 1414 AGTAGAATCAGCTCCCAATATACCGATCCATGAACTCTGTGTCGACGAGGTACTTACT 1463

Qy 3620 TGGACTAAAAGATCAAACTTCTTTTCAAACTCATTTTCTCATATAAAATATTGTACATTC 3679

Db 1464 ----- 1463

Qy 3680 GGGTCGAATTGCTTCCAAAGCTCTCTTCTCCAGAGAGGTATTGCGCCAAAGTTTACATA 3739

Db 1464 GGGTCGAATTGCTTCCAAAGCTCTCTTCTCCAGAGAGGTATTGCGCCAAAGTTTACATA 1523

Qy 3740 TCGAGAAGAACACAGAGAGGAGAACAAACAAACAGAACAAAGATATCCATGCGACT 3799

Db 1524 TCGAGAAGAACACAGAGAGGAGAACAAACAAACAGAACAAAGATATCCATGCGACT 1583

```

Qy 3800 TGATCTTAACTTCAGCTCAGTTAAACACCACTAGTTGATGATCAAGAGAGAGAGAAACAC 3859
Db 1584 TGATCTTAACTTCAGCTCAGTTAAACACCACTAGTTGATGATCAAGAGAGAGAGAAACAC 1643
Qy 3860 AGGATTTCTTGGAAATCGGATTAGATGCTTCAAGCTAATGATGATGAGAGAGAAACAGTTT 3919
Db 1644 AGGATTTCTTGGAAATCGGATTAGATGCTTCAAGCTAATGATGATGAGAGAGAAACAGTTT 1703
Qy 3920 TAAACCATACAAAAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAAACATCC 3979
Db 1704 TAAACCATACAAAAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAAACATCC 1763
Qy 3980 TATCATTCATGTGAACAGAAAGATCCCAACGGATCGGTTGGAAACTCAAGCTTCCAC 4039
Db 1764 TATCATTCATGTGAACAGAAAGATCCCAACGGATCGGTTGGAAACTCAAGCTTCCAC 1823
Qy 4040 ATGA 4043
Db 1824 ATGA 1827

RESULT 8
BT001096 1612 bp mRNA linear PLN 30-OCT-2002
LOCUS Arabidopsis thaliana clone C105127 (D) putative MYB-related
DEFINITION transcription factor CCAL (At2g46830) mRNA, complete cds.
ACCESSION BT001096
VERSION BT001096.1 GI:24429605
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1612)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,
Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,
Ecker, J.R. and Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1612)
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,
Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,
Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (30-OCT-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
Annotation based on July 2002 version of the Arabidopsis genome
submitted to Genbank.
FEATURES
source
    Location/Qualifiers
        1..1612
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            /chromosome="2"
            /clone="C105127 (D)"
            /note="This clone is in pUNI 51.
            ecotype: Columbia"
        1..1612
            /gene="At2g46830"
        1..1581
            /gene="At2g46830"
            /evidence=experimental
            /product="putative MYB-related transcription factor CCAL"
            /protein_id="AA061004.1"
            /db_xref="GI:24429606"
            /translations="MGALDIAIPPPRPKRPNPNYPKRTGSGTILMSKTVNDGKES
            LGSEKVSHPWANEDRQSKPEKTLQEDNCSDFTHQYLSAASSMNKSCIETSNAST

```

```

FREFLPSREBGSQNNRVRKESNLDNAKSLNENGNEQGPQTPMHIPLVLPLGSSITSS
LSHPSPDPSHPHTVAGDYQSFNHIINSTLLOTALYTAQTATFASFPDPSGSGSPVP
GNSPPNLAAMAAATAVAASAWRAANGLLPLCAPLSSGGFTSHPSSTGCPSCDVEYTKA
STLHGVSQREQEHSASKARSLSDESDVENKSKPVCHQEPSATPESDADKSGDAGD
RKQVDRSSCSNTPSSDDVEADASERQEDGTNGEVKXETNEDTNKPTQTSNARRSRI
SSNITDPKWSDEKRIAFQALFSREVLPOSFTYREHREHEEQOQOORYPMALDLNF
TAQUTPVDDDEKRENTGFLGLDASKLMSRGRGTGFKPYKRCSMEAKESRLNNPII
HVEQKDPKRMRLTQAST"
1582..1612
/gene="At2g46830"
BASE COUNT 505 a 366 c 381 t
ORIGIN
Query Match 30.9%; Score 1342; DB 8; Length 1612;
Best Local Similarity 87.0%; Pred. No. 6.3e-268;
Matches 1612; Conservative 0; Mismatches 0; Indels 240; Gaps 3;
Qy 2223 ATGGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCCTAAGCGTAAACCAACAAT 2282
Db 1 ATGGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCCTAAGCGTAAACCAACAAT 60
Qy 2283 CCTTATCTCGAAGACGGGAAGTGGAAAGATCCTTATGTCAAAAACGGGTGTGAATGAT 2342
Db 61 CCTTATCTCGAAGACGGGAAGTGGAAAGATCCTTATGTCAAAAACGGGTGTGAATGAT 120
Qy 2343 GGAAGAGAGTCCCTTGGATCAGAAAAAGTGTGGCATCTCTGAGGTGATTTTTCATGGTCATA 2402
Db 121 GGAAGAGAGTCCCTTGGATCAGAAAAAGTGTGGCATCTCTGAGGTGATTTTTCATGGTCATA 160
Qy 2403 TGGCATCTTTTTCAGTGTGTCACTTGTCTCTCATCTTATTAATACAGATTGTGTGCTT 2462
Db 161 TGGCATCTTTTTCAGTGTGTCACTTGTCTCTCATCTTATTAATACAGATTGTGTGCTT 160
Qy 2463 CGTTTATAGTGGCCATGAGATCGAACAAATCAAGCCTGAAGAGAAAACTCTGCAG 2522
Db 161 CGTTTATAGTGGCCATGAGATCGAACAAATCAAGCCTGAAGAGAAAACTCTGCAG 213
Qy 2523 GAAGACAACCTGTTTTCAGATTTGTTTTCACATCATCAGTATCTCTGCTGTCATCTCTCATGAAT 2582
Db 214 GAAGACAACCTGTTTTCAGATTTGTTTTCACATCATCAGTATCTCTGCTGTCATCTCTCATGAAT 273
Qy 2583 AAAAGTTGTATAGACATCAAAACGACATTTTCCGCGAGTCTTTCGCTTTCACGGGAA 2642
Db 274 AAAAGTTGTATAGACATCAAAACGACATTTTCCGCGAGTCTTTCGCTTTCACGGGAA 333
Qy 2643 GAGGTAAACCAACATCTTTTCATTGCTATTGAGGTGTTTAAACAGATTAGTACTTTCATG 2702
Db 334 GAGGTAAACCAACATCTTTTCATTGCTATTGAGGTGTTTAAACAGATTAGTACTTTCATG 334
Qy 2703 AAACATAAACCGTGGGGGAATAACAGCGGAAGTCAAGATAACAGGGTAAAGAGGAGTCAA 2762
Db 335 AAACATAAACCGTGGGGGAATAACAGCGGAAGTCAAGATAACAGGGTAAAGAGGAGTCAA 370
Qy 2763 ACTCAGATTGTAATGCAAAAATCTCTGGAATAACCGTAATGACAAAGGACCTCAGACTTATC 2822
Db 371 ACTCAGATTGTAATGCAAAAATCTCTGGAATAACCGTAATGACAAAGGACCTCAGACTTATC 430
Qy 2823 CGATGCATATCCCTGTGCTAGTGCCATTGGGAGCTCAATAACAGATTCTCTATCAGATC 2882
Db 431 CGATGCATATCCCTGTGCTAGTGCCATTGGGAGCTCAATAACAGATTCTCTATCAGATC 490
Qy 2883 CTCCTTCAGAGCCAGATAGTATCCCCACACAGTTGAGAGATTTATCAGTCGTTTTCCTA 2942
Db 491 CTCCTTCAGAGCCAGATAGTATCCCCACACAGTTGAGAGATTTATCAGTCGTTTTCCTA 550
Qy 2943 ATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTATCTGCGCAACTTTCGCT 3002
Db 551 ATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTATCTGCGCAACTTTCGCT 610
Qy 3003 CATCATTTTGGCCTCCCGATTCTTAGTGGTGGCTCACTGTTCAGGGAACCTCACCTCGA 3062
Db 611 CATCATTTTGGCCTCCCGATTCTTAGTGGTGGCTCACTGTTCAGGGAACCTCACCTCGA 670

```

```
QY 3063 ATCTGGCTGCCATGGCGGCGGCACTGTTGCAGCTGCTAGTGTCTTGGTGGCTGCGCAATG 3122
|
|
|
Db 671 ATCTGGCTGCCATGGCGGCGGCACTGTTGCAGCTGCTAGTGTCTTGGTGGCTGCGCAATG 730
|
|
|
QY 3123 GATTATTACCTTTATGTGCTCTCTTAGTTCAGGTGGTTCCTAGTTCATCTCCATCTA 3182
|
|
|
Db 731 GATTATTACCTTTATGTGCTCTCTTAGTTCAGGTGGTTCCTAGTTCATCTCCATCTA 790
|
|
|
QY 3183 CTTTGGACATCATGTGATGTAGTACACAAAAGCAAGCACTTTACAACTGGTTCTG 3242
|
|
|
Db 791 CTTTGGACATCATGTGATGTAGTACACAAAAGCAAGCACTTTACAACTGGTTCTG 850
|
|
|
QY 3243 TCGAGAGCGAGAGCAAGCACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTCAG 3302
|
|
|
Db 851 TCGAGAGCGAGAGCAAGCACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTCAG 910
|
|
|
QY 3303 AGGATGTTGAAATAAGAGTAACCACTTTGTTCATGAGCAGCGCTTCTGCAACACCTTGAGA 3362
|
|
|
Db 911 AGGATGTTGAAATAAGAGTAACCACTTTGTTCATGAGCAGCGCTTCTGCAACACCTTGAGA 970
|
|
|
QY 3363 GTGATCAAGGGTTTCAGATGGAGCAGGAGACAGAAAACAAAGTTGACCGGTCCTCGTGTG 3422
|
|
|
Db 971 GTGATCAAGGGTTTCAGATGGAGCAGGAGACAGAAAACAAAGTTGACCGGTCCTCGTGTG 1030
|
|
|
QY 3423 GCTCAAACTCCGTCGAGTGTGATGATGTTGAGGCGGATGCATCAGAAAGCAAGAGG 3482
|
|
|
Db 1031 GCTCAAACTCCGTCGAGTGTGATGATGTTGAGGCGGATGCATCAGAAAGCAAGAGG 1090
|
|
|
QY 3483 ATGGCAACCAATGTGAGGTGAAGAAACGAATGAAGACACTATAAACCTTCAAACTTCAG 3542
|
|
|
Db 1091 ATGGCAACCAATGTGAGGTGAAGAAACGAATGAAGACACTATAAACCTTCAAACTTCAG 1150
|
|
|
QY 3543 AGTCCAAATGCAGCCGAGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGT 3602
|
|
|
Db 1151 AGTCCAAATGCAGCCGAGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGT 1210
|
|
|
QY 3603 CTGACGAGGTACTTACTTGGACTAAGATCAACTTCTTTATTTCAAACTTTCTCAT 3662
|
|
|
Db 1211 CTGACGA-----GGTGCAGTTCCTTCAAGCTCTCTTCCAGAGAGGTATTG 1272
|
|
|
QY 3663 ATAATATTTACATTCGGGTGGAATTTGGCTTCCAAAGCTCTCTTCCAGAGAGGTATTG 3722
|
|
|
Db 1218 -----GGTGCAGTTCCTTCCAAAGCTCTCTTCCAGAGAGGTATTG 1260
|
|
|
QY 3723 CGCAGAAAGTTTACATATCGAGAGAAACACAGAGAGGAAGAACAAACAAAGAACAA 3782
|
|
|
Db 1261 CGCAGAAAGTTTACATATCGAGAGAAACACAGAGAGGAAGAACAAACAAAGAACAA 1320
|
|
|
QY 3783 AGATATCCAAATGCGACTTTCATCTTAATTTACAGCTCAGTTAACACAGTTGATGATCAA 3842
|
|
|
Db 1321 AGATATCCAAATGCGACTTTCATCTTAATTTACAGCTCAGTTAACACAGTTGATGATCAA 1380
|
|
|
QY 3843 GAGGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAGCTTAATGAGT 3902
|
|
|
Db 1381 GAGGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAAGCTTAATGAGT 1440
|
|
|
QY 3903 AGAGAGAGAAACAGGTTTAAACATACAAAGATGTTCCATGGAAGCCAAAGAAAGTAGA 3962
|
|
|
Db 1441 AGAGAGAGAAACAGGTTTAAACATACAAAGATGTTCCATGGAAGCCAAAGAAAGTAGA 1500
|
|
|
QY 3963 ATCTCTCAACAACTCTATCTTATGTTGGAACAGAAAGATCCCAAGCGATGCGGTTG 4022
|
|
|
Db 1501 ATCTCTCAACAACTCTATCTTATGTTGGAACAGAAAGATCCCAAGCGATGCGGTTG 1560
|
|
|
QY 4023 GAAACTCAAGCTTCCACATGAGACTCTATTTTTCATCTGATCTGTTGTTGTA 4074
|
|
|
Db 1561 GAAACTCAAGCTTCCACATGAGACTCTATTTTTCATCTGATCTGTTGTTGTA 1612
|
|
|
```

```
RESULT 9
BT001105 1612 bp mRNA linear PLN 07-NOV-2002
LOCUS Arabidopsis thaliana clone Cl05127 (D) putative MYB-related
DEFINITION
```

transcription factor CCA1 (At2g46830) mRNA, complete cds.  
BT001105  
BT001105.1 GI:24762204  
FLI CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

```
QY 2343 GGAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCTGAGTGATTTTTCATGGTCATA 2402
D 121 GGAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCTGAGTGATTTTTCATGGTCATA 160
QY 2403 TGGCATCTTTTGGCAGTGTGTACATGTCTCTCATGTATTAATACAGATTGTGTGCTT 2462
D 161 ----- 160
QY 2463 CGTTTATAGATGGCAATGAAGATCGACAAATCAAGCTCGAAGAGAAAACTCTGCGAG 2522
D 161 -----AGATGGCAATGAAGATCGACAAATCAAGCTCGAAGAGAAAACTCTGCGAG 213
QY 2523 GAAGACAATGTTTCAGATTGTTTCACTCATCAGTATCTCTCTGCTGCATCTCCATGAAT 2582
D 214 GAAGACAATGTTTCAGATTGTTTCACTCATCAGTATCTCTCTGCTGCATCTCCATGAAT 273
QY 2583 AAAAGTTGTATAGACATCAAGCGAAGCACTTTCCGCGAGTCTTTCGCTTCAAGGAA 2642
D 274 AAAAGTTGTATAGACATCAAGCGAAGCACTTTCCGCGAGTCTTTCGCTTCAAGGAA 333
QY 2643 GAGGTAAAAACAATCTTTTCATTTGCTATTTTGGAGTTTAAAGACGATTAGTACTTTTCATG 2702
D 334 G----- 334
QY 2703 AAATAAAACCGTGGGGAATAACAGGGAAGTCAGATAACAGGGTAAGAAAGGAGTCAA 2762
D 335 -----AGGGAAGTCAGATAACAGGGTAAGAAAGGAGTCAA 370
QY 2763 ACTCAGATTGAATGCAAAATCTCTGAAACCGTATGAGCAAGGACCTCAACATTATC 2822
D 371 ACTCAGATTGAATGCAAAATCTCTGAAACCGTATGAGCAAGGACCTCAACATTATC 430
QY 2823 CGATGATATCCCTGTGTAGTGCCATTGGGGAGCTCAATAACAAGTTCTCTATCACATC 2882
D 431 CGATGATATCCCTGTGTAGTGCCATTGGGGAGCTCAATAACAAGTTCTCTATCACATC 490
QY 2883 CTCCTTCAGAGCAGATAGTCATCCCAACAGTTGCGAGAGATTATCAGTCTGTTTCTTA 2942
D 491 CTCCTTCAGAGCAGATAGTCATCCCAACAGTTGCGAGAGATTATCAGTCTGTTTCTTA 550
QY 2943 ATCATATATGTCACCCCTTTTACAACACCGCTCTTTTACTGCGCAACTTTTCGCT 3002
D 551 ATCATATATGTCACCCCTTTTACAACACCGCTCTTTTACTGCGCAACTTTTCGCT 610
QY 3003 CATCAITTTGGCCTCCCGATTTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACTCCGA 3062
D 611 CATCAITTTGGCCTCCCGATTTAGTGGTGGCTCACCTGTTCCAGGGAACCTCACTCCGA 670
QY 3063 ATCTGGCTGCATGGCGGAGCACTGTTCCAGCTGTAGTGTGTTGGTGGCTGCGCAATG 3122
D 671 ATCTGGCTGCATGGCGGAGCACTGTTCCAGCTGTAGTGTGTTGGTGGCTGCGCAATG 730
QY 3123 GATTATTTACCTTTATGTGCTCTCTTAGTTTCAGTGGTGTTCCTAGTCTATCTCTCCATCTA 3182
D 731 GATTATTTACCTTTATGTGCTCTCTTAGTTTCAGTGGTGTTCCTAGTCTATCTCTCCATCTA 790
QY 3183 CTTTGGACCATCATGTGATGTAGATACACAAAGCAAGCACTTTTACAACTGGTTCTG 3242
D 791 CTTTGGACCATCATGTGATGTAGATACACAAAGCAAGCACTTTTACAACTGGTTCTG 850
QY 3243 TCGAGAGCCGAGAGCAAGACACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTCAG 3302
D 851 TCGAGAGCCGAGAGCAAGACACTCCGAGGCATCAAGGCTCGATCTTCACTGGACTCAG 910
QY 3303 AGGATGTTGAAATAAGATTAACACAGTTTGTCTAGAGCAGCTTCTGCAACACCTTGAGA 3362
D 911 AGGATGTTGAAATAAGATTAACACAGTTTGTCTAGAGCAGCTTCTGCAACACCTTGAGA 970
QY 3363 GTGATGCAAGGTTTCAGATGGAGCAGGACAGAAAAAAGTTGACCGGTCTCTGCTGTG 3422
D 971 GTGATGCAAGGTTTCAGATGGAGCAGGACAGAAAAAAGTTGACCGGTCTCTGCTGTG 1030
QY 3423 GCTCAACACTCGTTCGATGATGATGTTGAGCGGATGTCATCAGAAAGCAAGAGG 3482
```

```
D 1031 GCTCAAACTCCGTGAGTAGTGATGTTTTCAGGCGGATGTCATCAGAAAGCGCAAGG 1090
QY 3483 ATGCGACCAATGTGAGGTGAAAAGAAACGAATGAAGACACTATAAACCTTCAAACTTCAG 3542
D 1091 ATGCGACCAATGTGAGGTGAAAAGAAACGAATGAAGACACTATAAACCTTCAAACTTCAG 1150
QY 3543 AGTCCAAATGACCCCGCAGTAGAATCAGCTCCCAATATAACCGATCCATGAAGTCTGTGT 3602
D 1151 AGTCCAAATGACCCCGCAGTAGAATCAGCTCCCAATATAACCGATCCATGAAGTCTGTGT 1210
QY 3603 CTGACGAGTACTTACTTGGACTAAAGATCAACTTCTCTTTTATTTCAAATCATTTTCTCAT 3662
D 1211 CTGACGA----- 1217
QY 3663 ATAAATATTGTACATTGGGTGCAATTCCTTCCAGCTCTCTTCTCCAGAGAGTATTG 3722
D 1218 -----GGGTGCAATTCCTTCCAGCTCTCTTCTCCAGAGAGTATTG 1260
QY 3723 CCGCAAGTTTATACATATCGAGAAGAACACAGAGAGGAAGAACAAACAACAAGAACAA 3782
D 1261 CCGCAAGTTTATACATATCGAGAAGAACACAGAGAGGAAGAACAAACAACAAGAACAA 1320
QY 3783 AGATATCCAATGGCCTTGTATCTTAACTTCAAGCTCAGTTAAACACAGTTGATGATCAA 3842
D 1321 AGATATCCAATGGCCTTGTATCTTAACTTCAAGCTCAGTTAAACACAGTTGATGATCAA 1380
QY 3843 GAGGAGAGAGAAACACAGGATTTCTTGGNAATCGGATTAGATGCTTCAAAGCTTAATGAGT 3902
D 1381 GAGGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTAGATGCTTCAAAGCTTAATGAGT 1440
QY 3903 AGAGGAAGAACAGGTTTAAACCATACAAAGATGTTTCCATGGAAGCCAAAGAAAGTAGA 3962
D 1441 AGAGGAAGAACAGGTTTAAACCATACAAAGATGTTTCCATGGAAGCCAAAGAAAGTAGA 1500
QY 3963 ATCCTCAACAAACATCTATCATTTGATGTGGAACAGAAAGATCCCAACCGATGCGGTTG 4022
D 1501 ATCCTCAACAAACATCTATCATTTGATGTGGAACAGAAAGATCCCAACCGATGCGGTTG 1560
QY 4023 GAAACTCAAGCTTCCACATGAGACTATTTTTCATCTGATCTGTTTGTGA 4074
D 1561 GAAACTCAAGCTTCCACATGAGACTATTTTTCATCTGATCTGTTTGTGA 1612
```

## RESULT 10

AX509497 LOCUS AX509497 1134 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 4192 from Patent WO0216655.  
ACCESSION AX509497  
VERSION AX509497.1 GI:23390734

## KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

1 Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
Stress-regulated genes of plants, transgenic plants containing  
same, and methods of use  
Patent: WO 0216655-A 4192 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG  
(CH)

## FEATURES

source Location/Qualifiers

1..1134  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"

BASE COUNT 333 a 193 c 214 g 394 t

## ORIGIN

Query Match 26.1%; Score 1134; DB 6; Length 1134;  
Best Local Similarity 100.0%; Pred. No. 9.8e-225;



	Matches	1134;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	198	ATCAAAGGAGGAAGAAGAAGAGAGCCCTTTTGGAGCCATTTCATGAATTGGAATGAA	257							
Db	1	ATCAAAGGAGGAAGAAGAAGAAGAGCCCTTTTGGAGCCATTTCATGAATTGGAATGAA	60							
Qy	258	GGATATCAAAAAGATCTAACHACAAAGGCCAGCGCTTCTCCTTCAATCTTCTCTTGTA	317							
Db	61	GGATATCAAAAAGATCTAACHACAAAGGCCAGCGCTTCTCCTTCAATCTTCTCTTGTA	120							
Qy	318	CTAAATAATTTTCATCCCTTCTCTCTCTGTCTCTGTCTCTTCTTCTTCTTCTTCTTCT	377							
Db	121	CTAAATAATTTTCATCCCTTCTCTCTCTGTCTCTGTCTCTTCTTCTTCTTCTTCTTCT	180							
Qy	378	TCCATTTATGTCAAAAGTGTGTAAATTCCTCAAGACTATATAGAGATGTTTTGTTTCAT	437							
Db	181	TCCATTTATGTCAAAAGTGTGTAAATTCCTCAAGACTATATAGAGATGTTTTGTTTCAT	240							
Qy	438	TTTCCAAAATTTCAAACHTTTGTCCCCTTTTAGTCTTCTACCGTTCATGCATGTTAGCTT	497							
Db	241	TTTCCAAAATTTCAAACHTTTGTCCCCTTTTAGTCTTCTACCGTTCATGCATGTTAGCTT	300							
Qy	498	AGCTTAATGCTGAACGTGTTGAATACGATATGGCGCTTATGCTTAAAGAACAAAACTTA	557							
Db	301	AGCTTAATGCTGAACGTGTTGAATACGATATGGCGCTTATGCTTAAAGAACAAAACTTA	360							
Qy	558	TGGGTCTAAAAAAAATAAGCCCCAATATAAAACTATGGCCCCAAAATAAGTTAGTCCATT	617							
Db	361	TGGGTCTAAAAAAAATAAGCCCCAATATAAAACTATGGCCCCAAAATAAGTTAGTCCATT	420							
Qy	618	GAGTGTGAGATAGCGCGTGTAGTGAAACCGCACGAGAAATGCGGTTTCGATTGTTGGTGA	677							
Db	421	GAGTGTGAGATAGCGCGTGTAGTGAAACCGCACGAGAAATGCGGTTTCGATTGTTGGTGA	480							
Qy	678	GTAGTCGTCTAGATTCCTCGGGTCCACTCATCTTTCTAGTGTATCAGACACGTGTGCAAA	737							
Db	481	GTAGTCGTCTAGATTCCTCGGGTCCACTCATCTTTCTAGTGTATCAGACACGTGTGCAAA	540							
Qy	738	ACTGTGGGAGAGATTAAACGATCTTAACTAGGTCCCAGTCAAGATATTTATAACGAA	797							
Db	541	ACTGTGGGAGAGATTAAACGATCTTAACTAGGTCCCAGTCAAGATATTTATAACGAA	600							
Qy	798	TTGACCTTTTAAACCTTTTCAGTGTATCCCGAACTCGTGGCTTGAATACAAAGAGTT	857							
Db	601	TTGACCTTTTAAACCTTTTCAGTGTATCCCGAACTCGTGGCTTGAATACAAAGAGTT	660							
Qy	858	GTGAACAGTTGATCTTAAAGATGNCACAGATGTAACTTGAACAAAGCTGAATCATCTC	917							
Db	661	GTGAACAGTTGATCTTAAAGATGNCACAGATGTAACTTGAACAAAGCTGAATCATCTC	720							
Qy	918	TTGAGCCACTAGTATGTTGACATATGGCAGTTTCTTTTGTAGCCTCGAAATAAATAATT	977							
Db	721	TTGAGCCACTAGTATGTTGACATATGGCAGTTTCTTTTGTAGCCTCGAAATAAATAATT	780							
Qy	978	AAAAAGTTTGAGTTTAAAGATAATPATAGTGGCTGAGATTTCTCCATTTCOGTAGCTTCT	1037							
Db	781	AAAAAGTTTGAGTTTAAAGATAATPATAGTGGCTGAGATTTCTCCATTTCOGTAGCTTCT	840							
Qy	1038	GGTCTCTTTCTTTGTTTTCATGTGATCAAAAGCAATCACTTCTTCTTCTTCTTCTCTCG	1097							
Db	841	GGTCTCTTTCTTTGTTTTCATGTGATCAAAAGCAATCACTTCTTCTTCTTCTTCTCTCG	900							
Qy	1098	ATTTCTTACTGTGTTTCTTATCCACGMAATCTGGAAATFAAAAAATGGAATCTTTATCGAAT	1157							
Db	901	ATTTCTTACTGTGTTTCTTATCCACGMAATCTGGAAATFAAAAAATGGAATCTTTATCGAAT	960							
Qy	1158	CCAAGCTGATTTGTTTCTTTCATGTAATCATCTCTCTPAAAGGTACTTAAGATTGATTTA	1217							
Db	961	CCAAGCTGATTTGTTTCTTTCATGTAATCATCTCTCTPAAAGGTACTTAAGATTGATTTA	1020							
Qy	1218	TTGTTCATGGTCTTCTTATTTGTTTGAAGAATAAATTCATGATGATTTGTTTTGTTTGTG	1277							
Db	1021	TTGTTCATGGTCTTCTTATTTGTTTGAAGAATAAATTCATGATGATTTGTTTTGTTTGTG	1080							

Qy	1278	GATTAGTGGAAATTTTGTAAAGAGAAGATCTGAAGTTGTGTAGAGGAGCTTAGTG	1331
Db	1081	GATTAGTGGAAATTTTGTAAAGAGAAGATCTGAAGTTGTGTAGAGGAGCTTAGTG	1134
RESULT 11			
AY086711			
LOCUS	AY086711	895 bp	mrna linear PLN 14-APR-2003
DEFINITION	Arabidopsis thaliana clone 26967 mRNA, complete sequence.		
ACCESSION	AY086711		
VERSION	AY086711.1	GI:21405421	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 895) Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.		
TITLE	Full-length messenger RNA sequences greatly improve genome annotation		
JOURNAL	Genome Biol. 3 (6), RESEARCH0029 (2002)		
MEDLINE	22088475		
PUBMED	12093376		
REFERENCE	2 (bases 1 to 895)		
AUTHORS	Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
TITLE	Full-length cDNA from Arabidopsis thaliana		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 895)		
AUTHORS	Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA		
COMMENT	This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.		
FEATURES	Location/Qualifiers		
source	1..895		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/clone="26967"		
	126..650		
CDS	/codon_start=1		
	/product="unknown"		
	/protein_id="AA063765.1"		
	/db_xref="GI:21555057"		
	/translation="MASLSVSSSSSTIIDSRAPSLASASAPSPCISLIPTLIQSHHRAAKATAYCRKIVRVVTRATTEVGAPATTAEATTELPEIVKTAQAEWKVDKRYIGSLFAFVVAALGWSAGMISPIDRLPLVGVLELIGVGTGFTYKQLVFKPDREALF EKVKSTYKDIIGSS"		
BASE COUNT	224 a	208 c	175 g 288 t
ORIGIN			
Query Match	7.1%	Score 306.6;	DB 8; Length 895;
Best Local Similarity	95.7%	Pred. No. 5.2e-53;	



Matches	315;	Conservative	0;	Mismatches	14;	Indels	0;	Gaps	0;
Qy	114	AGGAATTTGTTGTTATCTGTTT	CAGGAGGCTTTGTTG	AAGAGGTC	AAAGACACATAC	173			
Db	567	AAGAACCTGGCTTTCAAAC	CAGACAGGAGGCTTTGTTG	AAGAGGTC	AAAGACACATAC	626			
Qy	174	AAAGACATATTTAGGAGCAGCTGAAT	CAAGGAGGAAGAA	GAGAGGAGGCTTTT	TTG	233			
Db	627	AAAGACATATTTAGGAGCAGCTGAAT	CAAGGAGGAAGAA	GAGAGGAGGCTTTT	TTG	686			
Qy	234	AGGCCATTTCATGAATTCGAAT	TGAAGGATATCAA	AAGAAATCTTAA	CACAAAGGCCACGTCCT	293			
Db	687	AGGCCATTTCATGAATTCGAAT	TGAAGGATATCAA	AAGAAATCTTAA	CACAAAGGCCACGTCCT	746			
Qy	294	TCCTTCAATCTTTCCCTTTCTGTA	CTAAATTTTCATCCTTT	CTCTCTCTCTGTCCTCT	353				
Db	747	TCCTTCAATCTTTCCCTTTCTGTA	CTAAATTTTCATCCTTT	CTCTCTCTCTGTCCTCT	806				
Qy	354	GGTCTTTTTAGCTCAAAGTATCAT	TCCATTTATGTCAAAGTGTTG	TAAATTCCTCAAGAC	413				
Db	807	GGTCTTTTTAGCTCAAAGTATCAT	TCCATTTATGTCAAAGTGTTG	TAAATTCCTCAAGAC	866				
Qy	414	TATATATGAGATGTTTTGTTTCAT	TTTTTCC	442					
Db	867	TATATATGAGATGTTTTGTTTCAT	TTTTTCC	895					

RESULT 12	AY065156	906 bp	mrna	linear	PLN 10-DEC-2001
LOCUS	AY065156				
DEFINITION	Arabidopsis thaliana unknown protein (At2g46820; F19D11.10) mRNA, complete cds.				
ACCESSION	AY065156				
VERSION	AY065156.1	GI:17473793			
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE	1 (bases 1 to 906)				
AUTHORS	Nguyen, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H., Cheuk, R., Chung, M. K., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S. X., Narusaka, M., Pham, P. K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinzaki, K., Ecker, J., Theologis, A. and Davis, R. W.				
TITLE	Direct Submission				
JOURNAL	Submitted (03-DEC-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
COMMENT	e-mail for correspondence: arabesequence.stanford.edu				

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shimozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, N., Palm, C. J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shinn, P., Yamada, K., Ecker, J. J., Theologis, A. and Davis, R. W.

Nguyen, M., (SSP/Stanford) and Seki, M., (RIKEN GSC) contributed equally to this work. Shinozaki, K., (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as PIs.

```

location/Qualifiers
1..906
source
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

```

[illegible]

```

/db_xref="taxon:3702"
/chromosome="II"
/clone="RAF06-08-111"
/notes="This clone is in pbluescript
ecotype: Columbia"
1..906
/gene="At2g46820; F19D11.10"
138..662
/gene="At2g46820; F19D11.10"
/codon_start=1
/product="unknown protein"
/protein_id="AAL38332.1"
/db_xref="GI:17473794"
/tranlation="MASLSVSSSTIIDSRAPPSRSLASASSPSCISLPTLP IQSH
RAAKATATYCRKIVRVNVTATTEVGEAPATTEAETTELPEIVKTAQEAWEKVVDDKYA
IGSLFAGVVAUWGSAGMISAIDRLPLVFGVLGVGTGFTYKSLVFKPDPREALF
EKVKSJYKDI LGS"
230 a 208 c 180 g 288 t
ch 7.0%; Score 305.6; DB 8; Length 906;
l Similarity 95.7%; Pred. No. 8.4e-53;
314; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
14 AGGAATATTGTTGTTATCTGTTTCAGGAGGCTTTGTTTCAGAAGGTCAAGACCATAC 173
79 AAGAACTGCTCTTCAACACAGACAGGAGCTTCTGTTGAGAAGGTCAAGACCATAC 638
74 AAGACATATTAGGAGCAGCTGAATCAAGGAGGAAGAAGAAGAGAGAGCTTTTGTG 233
39 AAGACATATTAGGAGCAGCTGAATCAAGGAGGAAGAAGAAGAAGAAGAGCTTTTGTG 698
34 AGGCCATTATGAATTTGGAAATGAAGATATCAAAAGAACTCAACACAAAGGCCACGTCCT 293
99 AGGCCATTATGAATTTGGAAATGAAGATATCAAAAGAACTCAACACAAAGGCCACGTCCT 758
94 TCCTTCAATCTTTCCCTTCTGTGTAACATAAATTTTCATCCTTTCTCTCTCTGTCCT 353
59 TCCTTCAATCTTTCCCTTCTGTGTAACATAAATTTTCATCCTTTCTCTCTCTCTCT 818
54 GGTCTTTTTTAGCTCAAAAGTATCATCCATTTATGTCAAAAGTGTGTAAATTCCTCAAGAC 413
19 GGTCTTTTTTAGCTCAAAAGTATCATCCATTTATGTCAAAAGTGTGTAAATTCCTCAAGAC 878
14 TATATATGAGATGTTTGTGTTTCAATTTTC 441
9 TATATATGAGATGTTTGTGTTTCAATTTTC 906

```

RESULT 13	
LOCUS	AY166504
DEFINITION	Brassica rapa subsp. pekinensis Myb-related transcription factor
ACCESSION	AY166504
VERSION	AY166504.1
KEYWORDS	GI:27357049
SOURCE	Brassica rapa subsp. pekinensis
ORGANISM	Brassica rapa subsp. pekinensis
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE	1 (bases 1 to 318) Jang,H. and Hur,Y.
AUTHORS	
TITLE	Study on the expression pattern of genes associated with flowering in chinese cabbage plants
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 318) Jang,H. and Hur,Y.
AUTHORS	
TITLE	Direct Submission
JOURNAL	Submitted (16-OCT-2002) Biology, Chungnam National University, Kungdong, Yuseong-ku, Deajeon 305-764, Korea
FEATURES	Location/Qualifiers

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEN (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Tripp, M., Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M. K., Kim, C., Lin, J., Liu, S. X., Pham, P. K., Sakano, H., Shim, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R. W.

Tripp, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as PIs.

Location/Qualifiers

1. 632

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="2"

/clone="U17260"

/note="This clone is in pENTR/SD-dTelo This is a cloned PCR product using RIKEN clone RAFL06-08-111 (AY065156) as a template

ecotype: Columbia"

1. 632

/gene="At2g46820"

/note="synonym: F19D11.10"

1. 525

/gene="At2g46820"

/codon\_start=1

/product="unknown protein"

/protein\_id="AA047902.1"

/db\_xref="GI:21386997"

/translation="MASLSVSSSTIIDSRAAPSSRLASASSPSCISLTPLPQSHSTRAAKATAYCRKIVRNVVTRATVEGPAATTEAETTELPEIVKTAQAEWVKDDKYAIGSLAFAGVWALGWSAGMISALDRLPLPGVLELVIGYTGWFTYKNLVFKPDREALF

161 a 151 c 147 g 173 t

3.9%; Score 168.6; DB 8; Length 632;

Similarity 92.7%; Pred No. 2.4e-24;

77; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

4 4 AGGAAATGTTGTTATCTGTTTCAGGAGGCTTTGTTTGAAAGGTCAAGAGCATAC 173

|||||

12 2 AAGAACCTGGTCTTCAAAACGACAGCAGGAGGCTTTGTTTGAAAGGTCAAGAGCATAC 501

|||||

4 4 AAGACATATTAGGAGCAGCTGAATCAAGGAGGAGAAAGAAAGAGAGCCTTTTGG 233

|||||

12 2 AAGACATATTAGGAGCAGCTGAATCAAGGAGGAGAAAGAAAGAGAGCCTTTTGG 561

|||||

4 4 AGGCATTCATGAATTGGAATGAAGGATATCAAAAGATCTAAACAAGGCCACGTCCT 293

|||||

12 2 AGGCATTCATGAATTGGAATGAAGGATATCAAAAGATCTAAACAAGGCCACGTCCT 621

|||||

4 4 TCCTTCAATCT 304

|||||

12 2 TCCTTCAATCT 632

ATH530832 126 bp DNA linear PLN 29-MAR-2003

Arabidopsis thaliana T-DNA flanking sequence, left border, clone 218A11.

AJ530832

AJ530832.1 GI:26799092

left border; T-DNA flanking sequence.

Arabidopsis thaliana (chale cross)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS

1  
Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., Lepiniec, L., Caboche, M., and Lecharny, A.

TITLE

T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites

JOURNAL  
MEDLINE  
PUBMED

EMBO Rep. 3 (12), 1152-1157 (2002)  
22363535  
1246565

REFERENCE  
AUTHORS

2 (bases 1 to 126)  
Balzerque, S.

JOURNAL

Direct Submission  
Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE

COMMENT

PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES

source

1..126  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/cultivar="Wassilewskija"  
/db\_xref="taxon:3702"  
/clone="218A11"  
/clone\_lib="Arabidopsis thaliana T-DNA insertion lines"

misc\_feature

1..126  
/note="T-DNA flanking sequence  
left border"  
47 a 15 c 38 g 26 t

BASE COUNT  
ORIGIN

Query Match 2.9%; Score 126; DB 8; Length 126;  
Best Local Similarity 100.0%; Pred. No. 2.2e-15;  
Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 TCAGGAGGCTTTGTTGAGAGGTCAGAGCACATACAAAGACATATTAGGAGCAGCT 195

Db 1 TCAGGAGGCTTTGTTGAGAGGTCAGAGCACATACAAAGACATATTAGGAGCAGCT 60

QY 196 GAATCAAGGAGGAGAGAGAGAGAGAGCCCTTTTGGAGCCATTTCATGAATTGGAATG 255

Db 61 GAATCAAGGAGGAGAGAGAGAGAGAGCCCTTTTGGAGCCATTTCATGAATTGGAATG 120

QY 256 AAGGAT 261

Db 121 AAGGAT 126

Search completed: November 24, 2003, 09:24:16  
Job time : 15542.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 18:46:12 ; Search time 1060.65 Seconds  
(without alignments)  
11055.799 Million cell updates/sec

Title: US-10-084-553-1

Perfect score: 4344

Sequence: 1 gcagtggtcacttacaaga.....actcataagtaagcacaaaa 4344

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_19Jun03:\*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4344	100.0	4344	19 AAV65381	Arabidopsis thalia
2	1527.6	35.2	2254	19 AAV65382	Arabidopsis thalia
3	1525.2	35.1	2240	22 AAD06664	A. thaliana trans
4	1525.2	35.1	2240	22 AAD05746	Arabidopsis thalia
5	1525.2	35.1	2240	22 AAD05773	Arabidopsis thalia
6	1349.2	31.1	1827	24 ABZ13695	Arabidopsis thalia
7	1134	26.1	1134	24 ABZ16387	Arabidopsis thalia
8	312.6	7.2	865	24 AEN98810	Arabidopsis thalia

9	306.6	7.1	846	21 AAC39815	Arabidopsis thalia
10	306.6	7.1	895	21 AAC34359	Arabidopsis thalia
11	304.6	7.0	844	21 AAC49446	Arabidopsis thalia
12	304.6	7.0	893	21 AAC47534	Arabidopsis thalia
13	121.2	2.8	507	21 AAC52051	Arabidopsis thalia
14	116.6	2.7	478	21 AAC41232	Arabidopsis thalia
15	94.4	2.2	2730	21 AAC47080	Arabidopsis thalia
16	90.4	2.1	666	19 AAV09298	Inverse PCR fragme
17	83.2	1.9	2526	19 AAV09296	Nucleotide sequenc
18	83.2	1.9	2526	22 AAD06648	A. thaliana trans
19	83.2	1.9	2526	22 AAD05761	Arabidopsis thalia
20	83.2	1.9	2526	22 AAD05787	Arabidopsis thalia
21	83.2	1.9	2526	24 ABK65208	Arabidopsis cDNA e
22	81.6	1.9	1938	24 ABZ13111	Arabidopsis thalia
23	66.8	1.5	592	25 ABX56693	Arabidopsis thalia
24	65.4	1.5	8524	22 AAS46588	Tumour suppressor
25	65.4	1.5	8524	24 ABLJ3841	Human immune syste
26	62.8	1.4	1148	21 AAC45095	Arabidopsis thalia
27	62.8	1.4	1330	21 AAC38730	Arabidopsis thalia
28	62.8	1.4	26997	22 AAS46747	Tumour suppressor
29	62.4	1.4	15373	24 ABL32466	Human immune syste
30	61.6	1.4	525	24 ABZ14536	Arabidopsis thalia
31	60	1.4	320	21 AAC56665	Eucalyptus grandis
32	59.8	1.4	17594	24 ABLJ4026	Human immune syste
33	59.6	1.4	12138	24 ABLJ3942	Human immune syste
34	59.6	1.4	12138	24 ABK28335	DNA transcription
35	59.4	1.4	6115	24 ABLJ3801	Human immune syste
36	59	1.4	5586	24 ABK40004	Human chemically p
37	58.8	1.4	40324	24 ABQ67150	Human angiogenesis
38	58.4	1.3	411	25 ABX49356	Bovine EST associa
39	57.6	1.3	8079	24 ABL92313	Chemically treated
40	57.4	1.3	6095	22 AAS46310	Tumour suppressor
41	57.4	1.3	6095	24 ABL70150	Chemically treated
42	57.4	1.3	6095	24 ABLJ2361	Human immune syste
43	57.4	1.3	6095	24 ABLJ4475	Human metastasis a
44	57.4	1.3	61020	22 AAS46788	Tumour suppressor
45	57	1.3	7455	24 ABLJ3759	Human immune syste

#### ALIGNMENTS

RESULT 1

AAV65381

ID AAV65381 standard; DNA; 4344 BP.

XX

AC AAV65381;

XX

DT 15-FEB-1999 (first entry)

XX

DE Arabidopsis phytochrome regulated transcription factor CCAL DNA.

XX

KW Phytochrome regulated transcription factor; CCAL;

KW chlorophyll binding protein; lhcb1\*3; circadian rhythm; flowering;

KW transgenic plant; ss.

XX

OS Arabidopsis thaliana ecotype Columbia.

XX

PH Key Location/Qualifiers

FT CDS 1332..4043

FT /tag= a

FT /note= "contains introns"

FT exon 1..1999

FT /tag= b

FT /number= 1

FT intron 1200..1283

FT /tag= c

FT /number= 1

FT exon 1332..1370

FT /tag= d

FT /number= 2

FT intron 1371..1448

FT /tag= e



Db 1021 CCAATTTCCGTAGCTCTGCTCTCTTTCTTTGTTTCATTTGATCAAAAGCAAAATCACTTCT 1080  
Qy 1081 TCTTCTTCTTCTCGATTTCTTACTGTTTCTTATCAACGAAATCTGGAATTAATAA 1140  
Db 1081 TCTTCTTCTTCTCGATTTCTTACTGTTTCTTATCAACGAAATCTGGAATTAATAA 1140  
Qy 1141 TGGAACTTTTATCGAAATCCAAGCTGATTTTGTGTTCTTCAATGGAATCATCTCTCTAAAG 1200  
Db 1141 TGGAACTTTTATCGAAATCCAAGCTGATTTTGTGTTCTTCAATGGAATCATCTCTCTAAAG 1200  
Qy 1201 TACTTAAGATTGATTTATTTGTCATGCTCTTCTTATTTGTTGATGAATCACTGACTTGA 1260  
Db 1201 TACTTAAGATTGATTTATTTGTCATGCTCTTCTTATTTGTTGATGAATCACTGACTTGA 1260  
Qy 1261 TTGTTTTTTTGTGTTGATTTAGTGGAAATTTTGAAGAGAAGATCTGAAGTTGTGTAGA 1320  
Db 1261 TTGTTTTTTTGTGTTGATTTAGTGGAAATTTTGTGAAGAGAAGATCTGAAGTTGTGTAGA 1320  
Qy 1321 GGAGCTTAGTGAGAGACAAATTCGTCTGGAGAAGATCTGGTTATTAAGGTAATAATAAC 1380  
Db 1321 GGAGCTTAGTGAGAGACAAATTCGTCTGGAGAAGATCTGGTTATTAAGGTAATAATAAC 1380  
Qy 1381 TAAATTTTAGGGGAAGATGATTTGTTTATAGGTGTCAGAAATTCAGAAATTTTAATGAACT 1440  
Db 1381 TAAATTTTAGGGGAAGATGATTTGTTTATAGGTGTCAGAAATTCAGAAATTTTAATGAACT 1440  
Qy 1441 TGATATAGACTCGGAAGCCATATACGATAACAAGCAACGTGAAAGGTGGACTGAGGAAG 1500  
Db 1441 TGATATAGACTCGGAAGCCATATACGATAACAAGCAACGTGAAAGGTGGACTGAGGAAG 1500  
Qy 1501 AACATAATAGATTCATTTGAAGCTTTGAGGCTTTATGTTAGAGCATGGCAGAAATTTGAAG 1560  
Db 1501 AACATAATAGATTCATTTGAAGCTTTGAGGCTTTATGTTAGAGCATGGCAGAAATTTGAAG 1560  
Qy 1561 GTTGATTTTATTTCCCTTTATATGTCCTTATTTGTTGTTGTCAGAGGTTTGTCTTCAA 1620  
Db 1561 GTTGATTTTATTTCCCTTTATATGTCCTTATTTGTTGTTGTCAGAGGTTTGTCTTCAA 1620  
Qy 1621 ACTGATTTGCTTTTTTTCATTTTGGACAGAACATGTAGCAACAAAACTGCTGCCAGATA 1680  
Db 1621 ACTGATTTGCTTTTTTTCATTTTGGACAGAACATGTAGCAACAAAACTGCTGCCAGATA 1680  
Qy 1681 AGAAGTCAGCTCAGAAATTTTCTCAAGGTAAATCGGTTAAATTTGAAATGATGTTTC 1740  
Db 1681 AGAAGTCAGCTCAGAAATTTTCTCAAGGTAAATCGGTTAAATTTGAAATGATGTTTC 1740  
Qy 1741 TCATCTTCAATGGCTTAATGCTTAAGACTTATTGAAAGCCAGCAAGTTTCTGCTTCTT 1800  
Db 1741 TCATCTTCAATGGCTTAATGCTTAAGACTTATTGAAAGCCAGCAAGTTTCTGCTTCTT 1800  
Qy 1801 TTGCTTCTTTAGTCAGGAGATAGATTAAGTTACGTTTTTATAGATTTAGTAATGAGCAATAAG 1860  
Db 1801 TTGCTTCTTTAGTCAGGAGATAGATTAAGTTACGTTTTTATAGATTTAGTAATGAGCAATAAG 1860  
Qy 1861 TCTTAAATAGTTGGAGAAATGACGAGATGAATCGTTTTCTTTGTTTATGCTTATATC 1920  
Db 1861 TCTTAAATAGTTGGAGAAATGACGAGATGAATCGTTTTCTTTGTTTATGCTTATATC 1920  
Qy 1921 TTGTTAATCCAAAAATGTACATAGATTTCTTCAAGAAATGTTAGTTTCTTTTAGATTCT 1980  
Db 1921 TTGTTAATCCAAAAATGTACATAGATTTCTTCAAGAAATGTTAGTTTCTTTTAGATTCT 1980  
Qy 1981 TCAGATAAATCTTGTCCTTTACCGATTTCTGAGGTAGTGGCAAAAGTGGGCTGAGTGCT 2040  
Db 1981 TCAGATAAATCTTGTCCTTTACCGATTTCTGAGGTAGTGGCAAAAGTGGGCTGAGTGCT 2040  
Qy 2041 AGAAATTTTGAATGTTCTCTGATAGCCATAGAGGTAAACCAATTTTGAATTTTCCAG 2100  
Db 2041 AGAAATTTTGAATGTTCTCTGATAGCCATAGAGGTAAACCAATTTTGAATTTTCCAG 2100  
Qy 2101 TTCTGTCAATTTAAACTTTGTTAGGTGTCATTAGATTTTGTGTTTACGTTTGTATTAGAG 2160  
Db 2101 TTCTGTCAATTTAAACTTTGTTAGGTGTCATTAGATTTTGTGTTTACGTTTGTATTAGAG 2160

Qy 2161 GGTAAACAAAATACTACTCTCATCTCTCAGGTAGAGAAAGGCTGAAGCTTAAAGGTGTAG 2220  
Db 2161 GGTAAACAAAATACTACTCTCATCTCTCAGGTAGAGAAAGGCTGAAGCTTAAAGGTGTAG 2220  
Qy 2221 CTATGGGTCAAGCGCTAGACATAGCTATTCTCTCCAGGCTTAAGCGTAAACCAAACA 2280  
Db 2221 CTATGGGTCAAGCGCTAGACATAGCTATTCTCTCCAGGCTTAAGCGTAAACCAAACA 2280  
Qy 2281 ATCCCTTATCCTCGAAGACGGGAAGTGGAAACGATCTTATGTCAAAACGGGTGTGAATG 2340  
Db 2281 ATCCCTTATCCTCGAAGACGGGAAGTGGAAACGATCTTATGTCAAAACGGGTGTGAATG 2340  
Qy 2341 ATGGAAGAGAGTCCCTTTGGATCAGAAAAAGTGTGCGATCTCTGAGGTGATTTTTCATGTGCA 2400  
Db 2341 ATGGAAGAGAGTCCCTTTGGATCAGAAAAAGTGTGCGATCTCTGAGGTGATTTTTCATGTGCA 2400  
Qy 2401 TATGGCATCTTTTTCAGTGTGTACATTTGCTCTCATGTTTATTAATACAGATTTGTGTGC 2460  
Db 2401 TATGGCATCTTTTTCAGTGTGTACATTTGCTCTCATGTTTATTAATACAGATTTGTGTGC 2460  
Qy 2461 TTGCGTTTATAGATGGCCAAATGAGATCGAACAAATCAAGCCCTGAGAGAAACTCTGC 2520  
Db 2461 TTGCGTTTATAGATGGCCAAATGAGATCGAACAAATCAAGCCCTGAGAGAAACTCTGC 2520  
Qy 2521 AGGAAGACAACCTGTTTTCAGATTTGTTTCACTCATCAGTATCTCTGCTGCACTCTCCATGA 2580  
Db 2521 AGGAAGACAACCTGTTTTCAGATTTGTTTCACTCATCAGTATCTCTGCTGCACTCTCCATGA 2580  
Qy 2581 ATAAAGTTGTATAGAGACATCAAAACGAGCACTTTCCGCGAGTTCTTTCCTTTCACGGG 2640  
Db 2581 ATAAAGTTGTATAGAGACATCAAAACGAGCACTTTCCGCGAGTTCTTTCCTTTCACGGG 2640  
Qy 2641 AAGAGTAAAAAACATCTTTCATTTGCTATTGAGGTTTTTAAGACATTTAGTACTTTTCA 2700  
Db 2641 AAGAGTAAAAAACATCTTTCATTTGCTATTGAGGTTTTTAAGACATTTAGTACTTTTCA 2700  
Qy 2701 TGAACATAAAAACCGTGGGGGAATTAACAGGGAAGTCAAGATAACAGGCTTAAGAAGGAGTC 2760  
Db 2701 TGAACATAAAAACCGTGGGGGAATTAACAGGGAAGTCAAGATAACAGGCTTAAGAAGGAGTC 2760  
Qy 2761 AAATCAGATTTGAAATGCAAAAATCTCTGAAAAACGGTAATGAGCAAGGACCTCAGACTTA 2820  
Db 2761 AAATCAGATTTGAAATGCAAAAATCTCTGAAAAACGGTAATGAGCAAGGACCTCAGACTTA 2820  
Qy 2821 TCCGATGCAATATCCCTGCTAGTGCATTTGGGAGCTCAATAACAAAGTTCTCTATCACA 2880  
Db 2821 TCCGATGCAATATCCCTGCTAGTGCATTTGGGAGCTCAATAACAAAGTTCTCTATCACA 2880  
Qy 2881 TCCTCTTTCAGAGCCAGATAGTCAATCCCAACACAGTTGAGGAGATTTATCAGTCGTTTCC 2940  
Db 2881 TCCTCTTTCAGAGCCAGATAGTCAATCCCAACACAGTTGAGGAGATTTATCAGTCGTTTCC 2940  
Qy 2941 TAAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTATCTGCGCAACTTTCGC 3000  
Db 2941 TAAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTATCTGCGCAACTTTCGC 3000  
Qy 3001 CTCATCATTTTGGCCTCCCGATTTAGTGTGCTCACCTGTTCAGGGAACCTCACTCC 3060  
Db 3001 CTCATCATTTTGGCCTCCCGATTTAGTGTGCTCACCTGTTCAGGGAACCTCACTCC 3060  
Qy 3061 GAAATCTGGCTGCATGGCCGAGCCACATCTGTCAGCTGCTAGTGTGTTGGTGGCTGCCAA 3120  
Db 3061 GAAATCTGGCTGCATGGCCGAGCCACATCTGTCAGCTGCTAGTGTGTTGGTGGCTGCCAA 3120  
Qy 3121 TGGATTTATTAATCTTTATGTCCTCTTCTTAGTTTCAAGTGGTTTCACTAGTCACTCTCCATC 3180  
Db 3121 TGGATTTATTAATCTTTATGTCCTCTTCTTAGTTTCAAGTGGTTTCACTAGTCACTCTCCATC 3180  
Qy 3181 TACTTTTGGACCATCATGTGATGTAGAGTACAAAAAGCAAGCACTTTTACACATGGTTTC 3240  
Db 3181 TACTTTTGGACCATCATGTGATGTAGAGTACAAAAAGCAAGCACTTTTACATGGTTTC 3240







DE A. thaliana transcription factor G680 homolog, G214 cDNA.

XX Plant transcription factor; phenotype; sugar sensing characteristic;  
 KW transgenic plant; plant yield; growth; germination; photosynthesis;  
 KW glyoxylate metabolism; respiration; pathogen response; wounding response;  
 KW cell cycle regulation; pigmentation; flowering; senescence; physiology;  
 KW storage organ; metabolism; ss.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers  
 PH 238..2064  
 FT /tag= a  
 FT /product= "Transcription factor homolog"

XX WO200135725-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-US31414.

XX 17-NOV-1999; 99US-0166228.

XX 17-APR-2000; 2000US-0197899.

XX 22-AUG-2000; 2000US-0227439.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX (JIAN/) JIANG C.

XX (HEAR/) HEARD J.

XX (PINE/) PINEDA O.

XX (PILG/) PILGRIM M.

XX (ADAM/) ADAM L.

XX (RIEC/) RIECHMANN J L.

XX (YUGG/) YU G.

XX (SAMA/) SAMAH R.

XX Jiang C, Heard J, Pineda O, Pilgrim M, Adam L, Riechmann JI;

XX Yu G, Samaha R;

XX WPI; 2001-335977/35.

XX P-PSDB; AA02563.

XX Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the sugar sensing characteristics of plants and increasing  
 PT yield, e.g. corn, potato and cotton plants -

XX Claim 4; Page 122-124; 15pp; English.

XX The patent relates to polynucleotides encoding 35 plant transcription  
 CC factors which may be used to modify phenotype associated with a plant's  
 CC sugar sensing characteristics and increasing yield when their expression  
 CC level is altered. Sugars are central regulatory molecules that control  
 CC aspects of physiology, metabolism and development. Therefore the cDNAs  
 CC and proteins of the invention are useful for modifying the growth and  
 CC germination rates of plants, photosynthesis, glyoxylate metabolism,  
 CC respiration, starch and sucrose synthesis and degradation, pathogen  
 CC response, wounding response, cell cycle regulation, pigmentation,  
 CC flowering and senescence of plants and for modifying sink-source  
 CC relationships in seeds, tubers, roots, and other storage organs leading  
 CC to an increase in yield. The transcription factor polynucleotides and  
 CC polypeptides may be used to alter the structure and developmental  
 CC characteristics of plants such as soybean, wheat, corn, potato, cotton,  
 CC rice, oilseed, sunflower, alfalfa, sugarcane, turf, banana, blackberry,  
 CC blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower,  
 CC coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon,  
 CC onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn,  
 CC tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas.  
 CC The present sequence is a homolog of Arabidopsis thaliana transcription  
 CC factor cDNA.

XX Sequence 2240 BP; 685 A; 473 C; 477 G; 605 T; 0 other;

XX Query Match

XX Best Local Similarity

XX 35.1%; Score 1525.2; DB 22; Length 2240;

XX 88.1%; Pred. No. 0;

		Matches 1797;	Conservative	0;	Mismatches	3;	Indels	240;	Gaps	3;
QY	2180	TCTCTCTCAGT	TAGAGAGAGGCT	TGAAGCT	TAAAGGT	GTAGCT	ATAGGCT	CAAGCCCT	TAGA	2239
DB	441	TTTCTCCAA	GGTAGAGAAAGGCT	TGAAGCT	TAAAGGT	GTAGCT	ATAGGCT	CAAGCCCT	TAGA	500
QY	2240	CATAGCTATT	CTCTCCACGCT	CTTAAGCGT	TAAACCAAC	CAATCT	TATCTCT	CGAAGAC		2299
DB	501	CATAGCTATT	CTCTCCACGCT	CTTAAGCGT	TAAACCAAC	CAATCT	TATCTCT	CGAAGAC		560
QY	2300	GGGAAGT	CGAAGCGAT	CTTTATGT	CAAAAA	CGGGT	GTGAAT	GATGAA	AGAGT	2359
DB	561	GGGAAGT	CGAAGCGAT	CTTTATGT	CAAAAA	CGGGT	GTGAAT	GATGAA	AGAGT	620
QY	2360	ATCAGAAAA	AGTGTGCGCAT	CTCTGAGGT	GATTTT	TCATG	GTTCAT	TGTCAT	TGTCAT	2419
DB	621	ATCAGAAAA	AGTGTGCGCAT	CTCTGAGGT	GATTTT	TCATG	GTTCAT	TGTCAT	TGTCAT	643
QY	2420	GTGTACAT	TGCTCTCTCAT	TGTTTAT	TATATAC	AGATT	TGTGTT	TCGTTT	TATAGT	2479
DB	644	GTGTACAT	TGCTCTCTCAT	TGTTTAT	TATATAC	AGATT	TGTGTT	TCGTTT	TATAGT	653
QY	2480	TGAAGAT	CGACAA	CNAATCA	AAAGCCT	TGAAGAG	AAAACT	CTGCA	GAGAA	2539
DB	654	TGAAGAT	CGACAA	CNAATCA	AAAGCCT	TGAAGAG	AAAACT	CTGCA	GAGAA	713
QY	2540	TTGTTT	CATCAT	CAGTAT	CTCTGCT	GCATCT	CCTCC	ATGA	TAAAGT	2599
DB	714	TTGTTT	CATCAT	CAGTAT	CTCTGCT	GCATCT	CCTCC	ATGA	TAAAGT	773
QY	2600	ATCAAA	CGCAAG	CACCTTT	CCGCG	AGTTCT	TTCGCTT	CACGGA	AGGTAA	2659
DB	774	ATCAAA	CGCAAG	CACCTTT	CCGCG	AGTTCT	TTCGCTT	CACGGA	AGGTAA	817
QY	2660	TTCA	TGCTATT	TGAGGT	TTTAA	GACGAT	TAGTACT	TTTTCAT	GAAACT	2719
DB	818	TTCA	TGCTATT	TGAGGT	TTTAA	GACGAT	TAGTACT	TTTTCAT	GAAACT	817
QY	2720	GAATAA	CAGGAA	AGT	CAGAA	TAAACAGG	GTAAAG	AGAGT	CAAACT	2779
DB	818	GAATAA	CAGGAA	AGT	CAGAA	TAAACAGG	GTAAAG	AGAGT	CAAACT	870
QY	2780	AAATCT	CTGAAAA	CGGTAA	TATGAG	CAAG	CACTT	ATCCG	ATATCC	2839
DB	871	AAATCT	CTGAAAA	CGGTAA	TATGAG	CAAG	CACTT	ATCCG	ATATCC	930
QY	2840	CTAGT	GCCATT	TGGGAG	CTCAAT	TAAAG	TTCTCT	TATCAC	ATCTCT	2899
DB	931	CTAGT	GCCATT	TGGGAG	CTCAAT	TAAAG	TTCTCT	TATCAC	ATCTCT	990
QY	2900	AGTCAT	CCCCAC	AGTTG	CAAGG	AGATT	TATCAG	TGCTTT	CTCTAT	2959
DB	991	AGTCAT	CCCCAC	AGTTG	CAAGG	AGATT	TATCAG	TGCTTT	CTCTAT	1050
QY	2960	CTTTT	TACAA	CACCG	CTTTT	TATAC	TCTCC	CAAACT	TTTCGCT	3019
DB	1051	CTTTT	TACAA	CACCG	CTTTT	TATAC	TCTCC	CAAACT	TTTCGCT	1110
QY	3020	GATTCT	AGTGT	GCTCAC	CTGTT	TCAGG	GAATCT	CACTCC	GAAATCT	3079
DB	1111	GATTCT	AGTGT	GCTCAC	CTGTT	TCAGG	GAATCT	CACTCC	GAAATCT	1170
QY	3080	GCAGC	CACTGT	TTCAG	CTGTAG	TGCTT	TGGTGG	CTGCC	ATGGAAT	3139
DB	1171	GCAGC	CACTGT	TTCAG	CTGTAG	TGCTT	TGGTGG	CTGCC	ATGGAAT	1230
QY	3140	GCTCCT	CTTAGT	TTCAG	TGTTT	CACTAG	TCTCCT	CACTAT	CTCTACT	3199
DB	1231	GCTCCT	CTTAGT	TTCAG	TGTTT	CACTAG	TCTCCT	CACTAT	CTCTACT	1290
QY	3200	GATGTAG	AGTAC	AAAGCA	AGC	ACTTT	TACAA	CATGGT	TCTGTG	3259
DB	1291	GATGTAG	AGTAC	AAAGCA	AGC	ACTTT	TACAA	CATGGT	TCTGTG	1350



Db 501 CATAGCTATTCTCTCCACGCCCTAAGCGTAACCAACAACTCTTATCTCTCGAAAGAC 560  
Qy 2300 GGGAGTGGAGCATCTTATGTCAAAACCGGTGTGAATGATGAAAGAGTCCCTTGG 2359  
Db 561 GGGAGTGGAGCATCTTATGTCAAAACCGGTGTGAATGATGAAAGAGTCCCTTGG 620  
Qy 2360 ATCAGAAAAGTGTCCGATCTCTGAGGTGATTTTCATGTGTATATGGCATCTTTTGGCAGT 2419  
Db 621 ATCAGAAAAGTGTCCGATCTCTG----- 643  
Qy 2420 GTGTACATTGCTCTCTCTATGTTTAAATACAGATTGTGTGCTTGGTTTATAGATGCCAA 2479  
Db 644 -----AGATGGCCAA 653  
Qy 2480 TGAAGATCGACAAACAATCAAGCCTGAAGAGAAAATCTCTGAGGAAGACAACTGTTTCA 2539  
Db 654 TGAAGATCGACAAACAATCAAGCCTGAAGAGAAAATCTCTGAGGAAGACAACTGTTTCA 713  
Qy 2540 TTGTTTCTACTCATAGTATCTCTCTGCTGCATCTCCATGAATAAAAGTTGTATAGAGAC 2599  
Db 714 TTGTTTCTACTCATAGTATCTCTCTGCTGCATCTCCATGAATAAAAGTTGTATAGAGAC 773  
Qy 2600 ATCAACCGCAGCATCTTCCGCGAGTCTTCCGCTTACGCGGAGAGGTAAAAACAATCT 2659  
Db 774 ATCAACCGCAGCATCTTCCGCGAGTCTTCCGCTTACGCGGAG----- 817  
Qy 2660 TTCTATTGCTATTGAGGTTTAAAGACGATTAGTACTTTTCTATGATAAACTAAACCCGTGGG 2719  
Db 818 ----- 817  
Qy 2720 GAATAACAGGGAAGTCAGAAATAACAGGGTAAGAAAGGAGTCAAACTCAGATTGTAATGCA 2779  
Db 818 -----AGGGAAGTCAGAAATAACAGGGTAAGAAAGGAGTCAAACTCAGATTGTAATGCA 870  
Qy 2780 AAATCTCTGGAAAACGGTAATCAGCAAGACCTCAGACTTATCCGATGCATATCCCTGTG 2839  
Db 871 AAATCTCTGGAAAACGGTAATCAGCAAGACCTCAGACTTATCCGATGCATATCCCTGTG 930  
Qy 2840 CTAGTGCCATTGGGAGCTCAATAACAAAGTCTCTATCATCTCTCTTCCAGAGCCAGAT 2899  
Db 931 CTAGTGCCATTGGGAGCTCAATAACAAAGTCTCTATCATCTCTCTTCCAGAGCCAGAT 990  
Qy 2900 AGTCATCCCCACACAGTTGCGAGGATATCAGTCTGTTTCTTAAATCATATATATGTCACCC 2959  
Db 991 AGTCATCCCCACACAGTTGCGAGGATATCAGTCTGTTTCTTAAATCATATATATGTCACCC 1050  
Qy 2960 CTTTTCACAAACACCGCTTATATCTGCGCAACTTTCGCTCATCATTTTGGCTCTCC 3019  
Db 1051 CTTTTCACAAACACCGCTTATATCTGCGCAACTTTCGCTCATCATTTTGGCTCTCC 1110  
Qy 3020 GATTCAGTGTGGCTCACCTGTTCCAGGAACTCACCTCCGAATCTGGCTGCCATGGCC 3079  
Db 1111 GATTCAGTGTGGCTCACCTGTTCCAGGAACTCACCTCCGAATCTGGCTGCCATGGCC 1170  
Qy 3080 GCAGCAGTGTGAGCTGTAGTCTGTTGGTGGCTGCGCAATGGAATATTAACCTTTATGT 3139  
Db 1171 GCAGCAGTGTGAGCTGTAGTCTGTTGGTGGCTGCGCAATGGAATATTAACCTTTATGT 1230  
Qy 3140 GCTCTCTTTAGTTTCAAGTGTGTTTCACTAGTCACTCTCCATCTTCTTTTGGACCATCATGT 3199  
Db 1231 GCTCTCTTTAGTTTCAAGTGTGTTTCACTAGTCACTCTCCATCTTCTTTTGGACCATCATGT 1290  
Qy 3200 GATGTAGAGTACAAAAAGCAGCACTTTTCAACATGTTTCTGTGCGAGCGCGAGAGCAA 3259  
Db 1291 GATGTAGAGTACAAAAAGCAGCACTTTTCAACATGTTTCTGTGCGAGCGCGAGAGCAA 1350  
Qy 3260 GAACACTCGAGGCAATCAAGCTCGATCTTCACTGGACTCAGAGGATGTTGAAATAAG 3319  
Db 1351 GAACACTCGAGGCAATCAAGCTCGATCTTCACTGGACTCAGAGGATGTTGAAATAAG 1410  
Qy 3320 AGTAAACAGTGTGTCAGAGCAGCTTCTGCAACACCTGAGAGTATGCAAGAGGTTCA 3379  
Db 1411 AGTAAACAGTGTGTCAGAGCAGCTTCTGCAACACCTGAGAGTATGCAAGAGGTTCA 1470

Qy 3380 GATGAGCAGGAGACAGAAAACAAGTTGACCGTCTCTGTTGGCTCAAAACACTCCGTGG 3439  
Db 1471 GATGAGCAGGAGACAGAAAACAAGTTGACCGTCTCTGTTGGCTCAAAACACTCCGTGG 1530  
Qy 3440 AGTAGTGATGATGTTTGGCGGATGCATCAGAAAGGCAAGAGATGGCAACCAATGTTGAG 3499  
Db 1531 AGTAGTGATGATGTTTGGCGGATGCATCAGAAAGGCAAGAGATGGCAACCAATGTTGAG 1590  
Qy 3500 GTGAAAGAAAAGTAAGAGACACTAATAAACTTCAAACTTTCAGAGTCCAATGACGCCCGC 3559  
Db 1591 GTGAAAGAAAAGTAAGAGACACTAATAAACTTCAAACTTTCAGAGTCCAATGACGCCCGC 1650  
Qy 3560 AGTACAATCAGCTCCCAATATATACCGATCCATCGAGAGTCTGTCTGACAGGTTACTTACT 3619  
Db 1651 AGTACAATCAGCTCCCAATATATACCGATCCATCGAGAGTCTGTCTGACAGGTTACTTACT 1700  
Qy 3620 TGGACTAAGATCAACTTCTTTTCAAAATCATTTTCTCATATAAAATATTGTATACATTC 3679  
Db 1701 ----- 1700  
Qy 3680 GGGTCGAATTGCTTCCAAAGCTCTCTTCTTCAGAGAGGTATTTGCCGCAAGTTTACATA 3739  
Db 1701 GGGTCGAATTGCTTCCAAAGCTCTCTTCTTCAGAGAGGTATTTGCCGCAAGTTTACATA 1760  
Qy 3740 TCGAGAAAGAACACAGAGAGGAGAACAAACAACAAGAAAGATATCCAATGGCACT 3799  
Db 1761 TCGAGAAAGAACACAGAGAGGAGAACAAACAACAAGAAAGATATCCAATGGCACT 1820  
Qy 3800 TGATCTTAACCTTTCAGAGCTCAGTTAACACACCTTTCATGATCAAGAGGAGAGAAACAC 3859  
Db 1821 TGATCTTAACCTTTCAGAGCTCAGTTAACACACCTTTCATGATCAAGAGGAGAGAAACAC 1880  
Qy 3860 AGGATTTCTTGGAAATCGGATTAGATGCTTCAAAGCTTAATGAGTAGAGGAGAAACAGTTT 3919  
Db 1881 AGGATTTCTTGGAAATCGGATTAGATGCTTCAAAGCTTAATGAGTAGAGGAGAAACAGTTT 1940  
Qy 3920 TAAACCATACAAAAGATGTTCCATCGAAGCCAAAGAAAGTAAATCCTCAACAAATCC 3979  
Db 1941 TAAACCATACAAAAGATGTTCCATCGAAGCCAAAGAAAGTAAATCCTCAACAAATCC 2000  
Qy 3980 TATCATTTGAGAAACAGAAAGATCCCAAAACGATGCGGTTGGAAACTCAAGCTTCCAC 4039  
Db 2001 TATCATTTGAGAAACAGAAAGATCCCAAAACGATGCGGTTGGAAACTCAAGCTTCCAC 2060  
Qy 4040 ATGAGACTCTATTTTTCATCTGATCTGTTTGTACTCTGTTTAAAGTTTTCAGAGCA 4099  
Db 2061 ATGAGACTCTATTTTTCATCTGATCTGTTTGTACTCTGTTTAAAGTTTTCAGAGCA 2120  
Qy 4100 CTGCTACATTTTCTTTTCTTTTGGAGGCTTGTATTTGTTTCTGTCATAGTCTTC 4159  
Db 2121 CTGCTACATTTTCTTTTCTTTTGGAGGCTTGTATTTGTTTCTGTCATAGTCTTC 2180  
Qy 4160 TGTAAACATTTGACTCTGATTTATTTCAACAAATCATAAACTGTTTAACTTTTTTTTCCA 4219  
Db 2181 TGTAAACATTTGACTCTGATTTATTTCAACAAATCATAAACTGTTTAACTTTTTTTTCCA 2240

## RESULT 5

AAD05773

ID AAD05773 standard; cDNA; 2240 BP.

XX AAD05773;

XX 31-JUL-2001 (first entry)

XX Arabidopsis thaliana transcription factor, G214 cDNA.

DE Transcription factor; biochemical characteristic; controlling element;  
KW structural characteristic; developmental characteristic; gene therapy;  
KW agricultural biotechnology; plant trait modification; ss.

XX Arabidopsis thaliana.

XX FH Key Location/Qualifiers  
 FT CDS 238..2064  
 FT /\*tag= a  
 FT /product= "Transcription factor, G214"  
 XX WO200136597-A1.  
 XX PN  
 XX PD  
 XX PF  
 XX PR 14-NOV-2000; 2000WO-US31344.  
 XX PR 17-NOV-1999; 99US-0166228.  
 XX PR 17-APR-2000; 2000US-0197899.  
 XX PR 22-AUG-2000; 2000US-0227439.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (CREE/) CREELMAN R.  
 PA (YUGG/) YU G.  
 PA (ADAM/) ADAM L.  
 PA (RIEC/) RIECHMANN J L.  
 PA (HEAR/) HEARD J.  
 PA (SAMA/) SAMAHA R.  
 PA (PILG/) PILGRIM M.  
 PA (PINE/) PINEDA O.  
 PA (JIANG/) JIANG C.  
 XX Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;  
 PI Pilgrim M, Pineda O, Jiang C;  
 XX WPI: 2001-335999/35.  
 DR P-PSDB; AAE01889.  
 XX Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the biochemical characteristics of plants e.g. corn,  
 PT potato and cotton plants -  
 XX Claim 4; Page 55-57; 127pp; English.  
 XX The present sequence is Arabidopsis thaliana transcription factor,  
 CC G214 cDNA. The transcription factor is used for altering a plant's  
 CC biochemical characteristics. The transcription factor may be used to  
 CC alter the structure and developmental characteristics of plants such as  
 CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,  
 CC alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry,  
 CC raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant,  
 CC grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,  
 CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,  
 CC rosaceous fruits and/or vegetable brassicas. Transcription factors are  
 CC key controlling elements of biological pathways and altering expression  
 CC levels of 1 or more transcription factors can change entire biological  
 CC pathways in an organism. Therefore manipulating transcription factor  
 CC levels in plants offers great potential in agricultural biotechnology  
 CC for modifying a plant's traits. Transcription factor cDNA is useful in  
 CC gene therapy.  
 XX Sequence 2240 BP; 685 A; 473 C; 477 G; 605 T; 0 other;  
 SQ  
 Query Match 35.1%; Score 1525.2; DB 22; Length 2240;  
 Best Local Similarity 88.1%; Pred. No. 0;  
 Matches 1797; Conservative 0; Mismatches 3; Indels 240; Gaps 3;  
 QY 2180 TCTCTCAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGTAGCTATGGGTCAAGCGCTAGA 2239  
 DB 441 TTTCTCAAAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGTAGCTATGGGTCAAGCGCTAGA 500  
 QY 2240 CATAGCTATCTCTCCACGGCTTAAGGTAACCAACAATCCTTATCTCTCGAAGAC 2299  
 DB 501 CATAGCTATCTCTCTCCACGGCTTAAGGTAACCAACAATCCTTATCTCTCGAAGAC 560  
 QY 2300 GCGAAGTGAACGATCCTTATGTCAAAAAACGGGTGTGAATGATGGAAGAGTCCCTCG 2359  
 DB 561 GCGAAGTGAACGATCCTTATGTCAAAAAACGGGTGTGAATGATGGAAGAGTCCCTCG 620

QY 2360 ATCAGAAAAAGTGTGCGATCCTGAGGTGATTTTTCATGTCATATGCGATCTTTTTCAGT 2419  
 DB 621 ATCAGAAAAAGTGTGCGATCCTG----- 643  
 QY 2420 GTGTCAATTCCTCCTCATGTTATTAAATACAGATTGTGTGCTTCGTTTATAGATGGCAA 2479  
 DB 644 -----AGATGGCAA 653  
 QY 2480 TGAAGATCGAACAAATCAAAGCCTGAAGAGAAAACTCTCAGGAAACACAACTCTTCA 2539  
 DB 654 TGAAGATCGAACAAATCAAAGCCTGAAGAGAAAACTCTCAGGAAACAACTCTTCA 713  
 QY 2540 TTGTTTCACTCATCAGTATCTCTCTGCTGCTCATCTCCATCAATAAAAAAGTTGTATAGAG 2599  
 DB 714 TTGTTTCACTCATCAGTATCTCTCTGCTGCTCATCTCCATCAATAAAAAAGTTGTATAGAG 773  
 QY 2600 ATCAAAACGGAAGCACCTTTCCGCGAGTCTTTGCGCTTCAACGGAAGAGGTAATAAACAATCT 2659  
 DB 774 ATCAAAACGGAAGCACCTTTCCGCGAGTCTTTGCGCTTCAACGGAAG----- 817  
 QY 2660 TTCAATTGCTATTGAGGTTTAAAGACGATTAGTACTTTTTCATGAAACTAAAAACCGTGGG 2719  
 DB 818 ----- 817  
 QY 2720 GAATAACAGGGAAGTCAAGGTAACAGGTAAGAAAGAGTCAAACTCAGATTTGAATGCA 2779  
 DB 818 -----AGGGAAGTCAAGGTAACAGGTAAGAAAGAGTCAAACTCAGATTTGAATGCA 870  
 QY 2780 AATCTCTGGAAAAACGGTAAATGAGCAAGACCTCAGACTTATCCGATGCATATCCCTGTG 2839  
 DB 871 AATCTCTGGAAAAACGGTAAATGAGCAAGACCTCAGACTTATCCGATGCATATCCCTGTG 930  
 QY 2840 CTAGTCCCATTTGGGAGCTCAATAACAAGTTCTCTATCACATCCTCTTCAGAGCCAGAT 2899  
 DB 931 CTAGTCCCATTTGGGAGCTCAATAACAAGTTCTCTATCACATCCTCTTCAGAGCCAGAT 990  
 QY 2900 AGTCATCCCCACACAGTTGAGGAGATTATCAGTCGTTTCTTAATCATATAATGTCAACC 2959  
 DB 991 AGTCATCCCCACACAGTTGAGGAGATTATCAGTCGTTTCTTAATCATATAATGTCAACC 1050  
 QY 2960 CTTTTCACAAACACCGGCTCTTTATATGCGCGAACTTTTGGCCTCATCATTTTGGCCTCCC 3019  
 DB 1051 CTTTTCACAAACACCGGCTCTTTATATGCGCGAACTTTTGGCCTCATCATTTTGGCCTCCC 1110  
 QY 3020 GATTTCTAGTGGTGGCTCACTGTTCAGGGAACCTACCTCCGAATCTGGCTGCCATGCC 3079  
 DB 1111 GATTTCTAGTGGTGGCTCACTGTTCAGGGAACCTACCTCCGAATCTGGCTGCCATGCC 1170  
 QY 3080 GCAGCCACTGTTGCAGCTGCTAGTGTGCTGGGCTGCCAATGGATTATTACCTTTATGT 3139  
 DB 1171 GCAGCCACTGTTGCAGCTGCTAGTGTGCTGGGCTGCCAATGGATTATTACCTTTATGT 1230  
 QY 3140 GCTCCTCTTAGTTTCAAGTGGTTCACATAGTCATCTCCATCTACTTTTGGACCATCATGT 3199  
 DB 1231 GCTCCTCTTAGTTTCAAGTGGTTCACATAGTCATCTCCATCTACTTTTGGACCATCATGT 1290  
 QY 3200 GATGTAGACTACAAAAACCAAGCACTTTTACAACATGGTTCTGTGAGAGCCGAGAGCAA 3259  
 DB 1291 GATGTAGACTACAAAAACCAAGCACTTTTACAACATGGTTCTGTGAGAGCCGAGAGCAA 1350  
 QY 3260 GAACACTCCGAGGCTCAAAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAAAATAG 3319  
 DB 1351 GAACACTCCGAGGCTCAAAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAAAATAG 1410  
 QY 3320 AGTAAACAGTTTGTATGAGGAGCCTTCTGCAACACCTTGAGAGTGTGATCAAAAGGTTCA 3379  
 DB 1411 AGTAAACAGTTTGTATGAGGAGCCTTCTGCAACACCTTGAGAGTGTGATCAAAAGGTTCA 1470  
 QY 3380 GATGAGAGGAGAGACAGAAAAACAAGTTGACCGGCTCTCTGTGGCTCAAAACACTCCGTCG 3439  
 DB 1471 GATGAGAGGAGAGACAGAAAAACAAGTTGACCGGCTCTCTGTGGCTCAAAACACTCCGTCG 1530



QY 2720 GAATAACAGGAGAGTACAGATAACAGAGGTAAGAAAGGAGTCAAACTCAGATTTGAATGCA 2779  
DB 581 -----AGGGAAGTCAGATAACAGAGGTAAGAAAGGAGTCAAACTCAGATTTGAATGCA 633  
QY 2780 AATCTCTGGAAAACGGTAATGACGAAGGACCTCAGACTTATCGGATGCATATCCCTGTG 2839  
DB 634 AATCTCTGGAAAACGGTAATGACGAAGGACCTCAGACTTATCGGATGCATATCCCTGTG 693  
QY 2840 CTAGTGCATTTGGGAGCTCAATAACAAAGTTCTCTATCACATCCTCTTCAGAGCCAGAT 2899  
DB 694 CTAGTGCATTTGGGAGCTCAATAACAAAGTTCTCTATCACATCCTCTTCAGAGCCAGAT 753  
QY 2900 AGTCATCCCCACACAGTTCGAGGAGATTAATCAGTCGTTTCCCTAATCATATAATGTCAACC 2959  
DB 754 AGTCATCCCCACACAGTTCGAGGAGATTAATCAGTCGTTTCCCTAATCATATAATGTCAACC 813  
QY 2960 CTTTTCACAAACACCGGCTCTTTTATCTGCGCGAACTTTTCGCTCATCATTTTGGCCTCCC 3019  
DB 814 CTTTTCACAAACACCGGCTCTTTTATCTGCGCGAACTTTTCGCTCATCATTTTGGCCTCCC 873  
QY 3020 GATTCTAGTGGTGGCTCACCTGTTTCCAGGGAATCACTCCGAATCTGGCTGCCATGGCC 3079  
DB 874 GATTCTAGTGGTGGCTCACCTGTTTCCAGGGAATCACTCCGAATCTGGCTGCCATGGCC 933  
QY 3080 GCAGCCACTGTTGCAGCTGCTAGTGTCTTGGTGGCTGCCAATGGAATATTACCTTTATGT 3139  
DB 934 GCAGCCACTGTTGCAGCTGCTAGTGTCTTGGTGGCTGCCAATGGAATATTACCTTTATGT 993  
QY 3140 GCTCCTCTTAGTTACGTTGTTTCACTAGTATCATCTCCATCTACTTTTGGACCATCATGT 3199  
DB 994 GCTCCTCTTAGTTACGTTGTTTCACTAGTATCATCTCCATCTACTTTTGGACCATCATGT 1053  
QY 3200 GATTGTAGTACACAAAAGCAAGCACTTTACACATGTTCTGTGCAGAGCCGAGAGCAA 3259  
DB 1054 GATTGTAGTACACAAAAGCAAGCACTTTACACATGTTCTGTGCAGAGCCGAGAGCAA 1113  
QY 3260 GAACACTCCGAGGCATCAAAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAAATAAG 3319  
DB 1114 GAACACTCCGAGGCATCAAAAGGCTCGATCTTCACTGGACTCAGAGGATGTTGAAAATAAG 1173  
QY 3320 AGTAAACCAAGTTGTCATGAGCAGCTTCTGCAACACCTGAGAGTGATGAAAGGTTCA 3379  
DB 1174 AGTAAACCAAGTTGTCATGAGCAGCTTCTGCAACACCTGAGAGTGATGAAAGGTTCA 1233  
QY 3380 GATGGAGCAGACAGAAAACAAAGTTGACCGGCTCTGCTGCTCAAAACACTCCGTCG 3439  
DB 1234 GATGGAGCAGACAGAAAACAAAGTTGACCGGCTCTGCTGCTCAAAACACTCCGTCG 1293  
QY 3440 AGTAGTGATGATGTTGAGGCGGATGCATCAGAAAGCAAGGATGSCACCAATGTTGAG 3499  
DB 1294 AGTAGTGATGATGTTGAGGCGGATGCATCAGAAAGCAAGGATGSCACCAATGTTGAG 1353  
QY 3500 GTCAGAAACGAATGAGACACTTAATAACCTCAAACTTCAGAGTCCAATGACGCCGC 3559  
DB 1354 GTCAGAAACGAATGAGACACTTAATAACCTCAAACTTCAGAGTCCAATGACGCCGC 1413  
QY 3560 AGTAGAATCAGCTCCCAATATAACCGATCCCATGGAAGTCTGTCTGACGAGGTACTTACT 3619  
DB 1414 AGTAGAATCAGCTCCCAATATAACCGATCCCATGGAAGTCTGTCTGACGA----- 1463  
QY 3620 TGGACTAAAGATCAACTCTCTTTATTTCAAATCATTTTCTCATATAAATTTGTACATTC 3679  
DB 1464 ----- 1463  
QY 3680 GGGTCGAATGCTCCCAAGCTCTCTCTCCAGAGGATTTCCCGCAAGTTTACATA 3739  
DB 1464 GGGTCGAATGCTCCCAAGCTCTCTCTCCAGAGGATTTCCCGCAAGTTTACATA 1523  
QY 3740 TCGAAGAACACAGAGAGGAAGAACAAACAAACAAAGAACAAAGATATCCAATGGCACT 3799  
DB 1524 TCGAAGAACACAGAGAGGAAGAACAAACAAACAAAGAACAAAGATATCCAATGGCACT 1583

QY 3800 TGATCTTAACCTTCACAGCTCAGTTAAACACACAGTTGATGATCAAGAGGAGAGAGAAACAC 3859  
DB 1584 TGATCTTAACCTTCACAGCTCAGTTAAACACACAGTTGATGATCAAGAGGAGAGAGAAACAC 1643  
QY 3860 AGGATTTCTTGGAAATCGGATTAGATGCTTCAAAAGCTTAATGAGTAGAGAGAAACAGGTTT 3919  
DB 1644 AGGATTTCTTGGAAATCGGATTAGATGCTTCAAAAGCTTAATGAGTAGAGAGAAACAGGTTT 1703  
QY 3920 TAAACCATACAAAGAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCTTCAACAAACATCC 3979  
DB 1704 TAAACCATACAAAGAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCTTCAACAAACATCC 1763  
QY 3980 TATCATTCATGTGGAACAGAAAGATCCCAAAACGAGTCGGTTGGAAACTCAAGCTTCCAC 4039  
DB 1764 TATCATTCATGTGGAACAGAAAGATCCCAAAACGAGTCGGTTGGAAACTCAAGCTTCCAC 1823  
QY 4040 ATGA 4043  
DB 1824 ATGA 1827  
RESULT 7  
ABZ16387  
ID ABZ16387 standard; DNA; 1134 BP.  
XX  
AC ABZ16387;  
XX  
AC 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 4192.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
28-FEB-2002.  
XX  
24-AUG-2001; 2001WO-US26685.  
XX  
24-AUG-2000; 2000US-227866P.  
PR 26-JAN-2001; 2001US-264647P.  
PR 22-JUN-2001; 2001US-300111P.  
XX  
(SCRI ) SCRIPPS RES INST.  
(SYGN ) SYNGENTA PARTICIPATIONS AG.  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX  
WPI; 2002-304127/34.  
XX  
PT Identifying a stress condition to which a plant cell has been exposed  
and producing plants with increased tolerance to these abiotic stresses  
XX  
PS Claim 144; SEQ ID NO 4192; 577pp + Sequence Listing; English.  
XX  
XX The invention relates to identifying a stress condition to which a plant  
cell has been exposed, comprising:  
CC (a) contacting nucleic acid representative of expressed polynucleotides  
in the plant cell with an array or probes representative of the plant  
cell genome; and  
CC (b) detecting a profile of expressed polynucleotides in the plant cell  
characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
with increased tolerance to abiotic stress. The present sequence is that  
of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
in methods of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.  
XX









RESULT 10  
AAC34359  
ID AAC34359 standard; DNA; 895 BP.  
XX AC AAC34359;  
XX AC AAC34359;  
DT 17-OCT-2000 (first entry)  
XX XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 6367.  
XX XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX XX  
OS Arabidopsis thaliana.  
XX XX  
PN EP1033405-A2.  
XX XX  
PD 06-SEP-2000.  
XX XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 18-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

	Query Match	7.1%;	Score 306.6;	DB 21;	Length 895;
	Best Local Similarity	95.7%;	Pred. No. 2.8e-63;		
	Matches 315;	Conservative 0;	Mismatches 14;	Indels 0;	Gaps 0
QY	114	AGGAAATTGTTTGTATCTGTTTCAGGGAGGGCTTTGTTGAGAAGGTCGAAGGACACATAC	173		
Db	567	AAGAACCTGGTCTTCAAACACGACAGGGAGGGCTTTGTTTGAGAAGGTCGAAGGACACATAC	626		
QY	174	AAAGACATATTAGGAGCAGCTGAATCAAAGGAGGAAGAAGAAAGAGAGAGCCCTTTTGTG	233		
Db	627	AAAGACATATTAGGAGCAGCTGAATCAAAGGAGGAAGAAGAAGAGAGAGCCCTTTTGTG	686		
QY	234	AGGCCATTCAATTTGAAATGAAGGATATCAAAAGAATCTAAACAAAGGCCACGTCTCT	293		
Db	687	AGGCCATTCAATTTGAAATGAAGGATATCAAAAGAATCTAAACAAAGGCCACGTCTCT	746		

[illegible]

PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147302.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149168.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 7.0%; Score 304.6; DB 21; Length 844;  
Best Local Similarity 95.7%; Pred. No. 8.4e-63;

```
Matches 313; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 114 AGGAAATGTTTCTTATCTGTTTCAGGAGGCTTTTGTGAGAAAGTCAAGAGCACATAC 173
    |||||
Db 518 AGAACCTGGTCTTCAACACAGACAGGAGGCTTTGTTTGAGAAAGTCAAGAGCACATAC 577
    |||||
QY 174 AAGACATATTAGGAGGAGCTGAATCAAGAGGAGGAAGAAAGAGAGAGCCTTTTGG 233
    |||||
Db 578 AAGACATATTAGGAGGAGCTGAATCAAGAGGAGGAAGAAAGAGAGAGCCTTTTGG 637
    |||||
QY 234 AGGCCATTGATGAATGGAATGAAGATATCAAAAGAAATCTACACAAAGGCCACGTCCT 293
    |||||
Db 638 AGGCCATTGATGAATGGAATGAAGATATCAAAAGAAATCTACACAAAGGCCACGTCCT 697
    |||||
QY 294 TCCTTCAATCTTCCCTTCTTGTAACTAAATAATTTTCATCCTTCTCTCTCTCTCTCTCT 353
    |||||
Db 698 TCCTTCAATCTTCCCTTCTTGTAACTAAATAATTTTCATCCTTCTCTCTCTCTCTCTCT 757
    |||||
QY 354 GGTCTTTTGTAGCTCAAGTATCATCATTTATGTCAAAGTGTGTAATTCCTCAAGAC 413
    |||||
Db 758 GGTCTTTTGTAGCTCAAGTATCATCATTTATGTCAAAGTGTGTAATTCCTCAAGAC 817
    |||||
QY 414 TATATATGAGATGTTTGTGTTTTCATTTT 440
    |||||
Db 818 TATATATGAGATGTTTGTGTTTTCATTTT 844
    |||||

RESULT 12
AAC47534
ID AAC47534 standard; DNA; 893 BP.
XX
AC AAC47534;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 54175.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
```

```

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 7.0%; Score 304.6; DB 21; Length 893;
Best Local Similarity 95.7%; Pred. No. 8.6e-63;
Matches 313; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 114 AGGAAATTGTTTGTATCTGTTTCAGGAGGCTTTGTTGAGAAGGTCAAGAGCACATAC 173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 174 AAAGACATATTAGGAGCAGCTGAATCAAAAGGAGGAAGAAGAAGAAGAGCCCTTTTGG 233
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 627 AAAGACATATTAGGAGCAGCTGAATCAAAAGGAGGAAGAAGAAGAAGAGCCCTTTTGG 686
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 234 AGCCATTCAATGGAATGAAGGATATCAAAAGAATCTTAACAAAAGGCCAGCTCT 293
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 687 AGCCATTCAATGGAATGAAGGATATCAAAAGAATCTTAACAAAAGGCCAGCTCT 746
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 294 TCCTTCATCTTCCCTTCTTGTAACTAAATAATTTTCATCCTTCTCTCTCTCTCTCT 353
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 747 TCCTTCATCTTCCCTTCTTGTAACTAAATAATTTTCATCCTTCTCTCTCTCTCTCT 806
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 354 GGTCTTTTTCAGCTCAAAGTATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGAC 413
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 807 GGTCTTTTTCAGCTCAAAGTATCATCCATTTATGTCAAAGTGTGTAATTCCTCAAGAC 866
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 414 TATATATGAGATGTTTGTTCATTTT 440
Db TATATATGAGATGTTTGTTCATTTT 893

RESULT 13
AAC52051
ID AAC52051 standard; DNA; 507 BP.
XX
AC AAC52051;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 70280.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

```

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131149.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143342.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151103.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 08-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.





```
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.7%; Score 116.6; DB 21; Length 478;
Best Local Similarity 89.9%; Pred. No. 7.9e-18;
Matches 125; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 114 AGGAAATGTTGTTATCTGTTTCAGGAGGCTTTTGTGAGAGGTTCAAGAGCATAAC 173
    |||||
Db 329 AAGAACCTGCTCTTCAACAGACAGGAGGCTTTGTTGAGAGGTTCAAGAGCATAAC 388
    |||||

QY 174 AAGACATATTAGGAGCAGCTGAATCAAGGAGGAGGAAGAAGAAGAGAGGCTTTTGTG 233
    |||||
Db 389 AAGACATATTAGGAGCAGCTGAATCAAGGAGGAGGAAGAAGAAGAAGAGAGGCTTTTGTG 448
    |||||

QY 234 AGGCCATTCAATGAATTGGA 252
    |||||
Db 449 AGGCCATTCAATGAATTGGA 467

RESULT 15
AAC47080
ID AAC47080 standard; DNA; 2730 BP.
XX
AC AAC47080;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 52490.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
```

PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.

```
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      2.2%; Score 94.4; DB 21; Length 2730;
Best Local Similarity 64.7%; Pred. No. 3.5e-12;
Matches 172; Conservative 0; Mismatches 66; Indels 28; Gaps 1;

QY 1447 AGACTCGGAGCCATATACGATAACAAAGCAACGCGAAGGTGGACTGAGGAAGAACATA 1506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 773 AGCAAGAAAGCCATATACAATAACAAAGCAGCGAGCGATGGACTGAGGATGAGCATG 832

QY 1507 ATAGATTTCATTGAAGCTTTGAGGCTTTATGGTAGACATGGCAGAGATTGAAGGTTGAT 1566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 833 AGAGGTTTCTAGAGGCTTTGAGGCTTTATGGAAAGACTTGGCAACGAATTTGAAGGAGTGT 892

QY 1567 TTTTATTTCCCTTTATATGCTTTATTTTGTGTTTGCAGAGGTTTGTCTTCAAACTGAT 1626
    |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 893 CTTTATCTCTCTTGATTGGGAGATT-----CCTTCT 924

QY 1627 TTGCTTTTTCATTTGGACAGACATGTAGCAACAAAACTGCTGTCAGATAAGAAGT 1686
    || |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 925 TTTCTTTTCATTTGTGTGCAAGACATATTTGGACAAAGACTGCTGTCAGATCAGAAGT 984

QY 1687 CACGCTCAGAAATTTTCTCCAAGT 1712
    || || || || || || || || || || || || || || || || || || || || || ||
Db 985 CATGCACAAAAGTTCTTCACAAAGTT 1010
```

Search completed: November 24, 2003, 02:51:10  
Job time : 1076.65 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 18:46:12 ; Search time 550.348 Seconds  
(without alignments)  
11055.799 Million cell updates/sec

Title: US-10-084-553-3

Perfect score: 2254

Sequence: 1 ttagatttccattccgt.....tttccaaaaa..... 2254

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2254	100.0	2254	19 AAV65382	Arabidopsis phyto
2	2240	99.4	2240	22 AAD06664	A. thaliana transc
3	2240	99.4	2240	22 AAD05746	Arabidopsis thalia
4	2240	99.4	2240	22 AAD05773	Arabidopsis thalia
5	1827	81.1	1827	24 ABZ13695	Arabidopsis thalia
6	1527.6	67.8	4344	19 AAV65381	Arabidopsis phyto
7	217.2	9.6	2526	19 AAV09296	Nucleotide sequenc
8	217.2	9.6	2526	22 AAD06648	A. thaliana transc

9	217.2	9.6	2526	22 AAD05761	Arabidopsis thalia
10	217.2	9.6	2526	22 AAD05787	Arabidopsis thalia
11	217.2	9.6	2526	24 ABK65208	Arabidopsis cDNA e
12	213	9.4	1938	24 ABZ13311	Arabidopsis thalia
13	189	8.4	1134	24 ABZ16387	Arabidopsis thalia
14	144.2	6.4	2730	21 AAC47080	Arabidopsis thalia
15	118.8	5.3	364	21 AAC57320	Pinus radiata tran
16	116.8	5.2	1164	24 ABK65272	Arabidopsis cDNA e
17	104.2	4.6	1148	21 ABK45095	Arabidopsis thalia
18	103.2	4.6	1199	21 AAC49487	Arabidopsis thalia
19	103	4.6	1166	22 AAD05830	Arabidopsis thalia
20	103	4.6	1197	21 AAC38743	Arabidopsis thalia
21	98.8	4.4	1391	22 AAD06644	A. thaliana transc
22	97.2	4.3	1416	21 AAC48314	Arabidopsis thalia
23	97.2	4.3	1419	21 AAC36749	Arabidopsis thalia
24	92.4	4.1	562	21 AAC56830	Pinus radiata tran
25	92.4	4.1	919	21 AAC56354	Pinus radiata tran
26	92.2	4.1	1330	21 AAC38730	Arabidopsis thalia
27	87.6	3.9	450	21 AAC57286	Eucalyptus grandis
28	87.6	3.9	541	21 AAC56124	Eucalyptus grandis
29	83.2	3.7	1612	21 AAC57364	Eucalyptus grandis
30	83	3.7	1237	22 AAD06660	A. thaliana transc
31	83	3.7	1352	21 AAC40200	Arabidopsis thalia
32	75.4	3.3	437	21 AAC57214	Eucalyptus grandis
33	75.4	3.3	666	19 AAV09298	Inverse PCR fragme
34	74.8	3.3	365	25 ABX18648	Human Gbp-mannose
35	73.8	3.3	373	21 AAC56764	Eucalyptus grandis
36	73	3.2	592	25 ABX56693	Arabidopsis thalia
37	68	3.0	320	21 AAC56665	Eucalyptus grandis
38	64	2.8	633	22 AAH87727	Eucalyptus grandis
39	53	2.4	397	25 ABX48619	Peppermint plant o
40	52.6	2.3	629	25 ABT21705	Bovine EST associa
41	50.6	2.2	375	23 ABV4911	Breast cancer mark
42	50.6	2.2	416	25 ABX39419	Human prostate exp
43	50.2	2.2	8910	24 ABN80133	Bovine EST associa
44	50	2.2	374	25 ABX47736	Human chemically m
45	50	2.2	410	22 AAI83503	Human polynucleoti

#### ALIGNMENTS

#### RESULT 1

AAV65382

ID AAV65382 standard; cDNA to mRNA; 2254 BP.

XX AC

XX AC

XX AC

DT 15-FEB-1999 (first entry)

XX AC

DE Arabidopsis phytochrome regulated transcription factor CCAL cDNA.

XX AC

KW Phytochrome regulated transcription factor; CCAL;

KW chlorophyll binding protein; Lhcb1\*3; circadian rhythm; flowering;

KW transgenic plant; ss.

XX AC

OS Arabidopsis thaliana ecotype Columbia.

XX AC

PH Key

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

FT CDS

DR P-PSDB; AAW79280.  
 XX Alteration of plant circadian rhythms - by transformation with a  
 PT transcription factor for the promoter of the Arabidopsis chlorophyll  
 PT binding protein  
 XX  
 PS Claim 1; Page 43-47; 47pp; English.  
 XX  
 CC This cDNA sequence codes for a phytochrome-regulated transcription  
 CC factor, designated CCA1 (see AAW79280), that binds to the promoter  
 CC region of the chlorophyll binding protein gene (Lhcb1\*3) of  
 CC Arabidopsis thaliana. To isolate the clone, a directional cDNA  
 CC expression library was constructed in lambda gt22A using cDNA  
 CC derived from leaves of Arabidopsis that had been grown in  
 CC continuous white light for 3 weeks. The library was screened with  
 CC an A2 fragment of the Lhcb1\*3 promoter. Overlapping phage clones  
 CC provided a partial sequence, and the entire 5' end was determined  
 CC by primer extension analysis. The CCA1 cDNA was subsequently used  
 CC to identify a genomic clone (see AAW65381). A claimed method for  
 CC altering plant response to daylength comprises transforming a plant  
 CC with either the CCA1 genomic DNA or cDNA, a hybridising sequence or  
 CC a nucleic acid encoding a protein containing amino acids 24-75 of  
 CC the 608-amino acid CCA1 protein. Manipulation of plant circadian  
 CC rhythms (e.g. by overexpression of CCA1) can be used to delay  
 CC flowering, and thus increase yield from both forage crops, where  
 CC biomass will be produced for longer, and from seed crops such as  
 CC rape, where a longer time in vegetative state will lead to larger  
 CC plants and thus a higher seed set (claimed).  
 XX  
 SQ Sequence 2254 BP; 699 A; 473 C; 477 G; 605 T; 0 other;

Query Match 100.0%; Score 2254; DB 19; Length 2254;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGATTTCCTCCATTCGAGTCTTCGGTCTCTTTCTTTCTTTCATTCATCAAGACA 60  
 DB 1 TGAGATTTCCTCCATTCGAGTCTTCGGTCTCTTTCTTTCTTTCATTCATCAAGACA 60

QY 61 AATCAGTCT 120  
 DB 61 AATCAGTCT 120

QY 121 GAATTAATAATGGAATCTTTATCGAATCCAGCTGATTTTGTCTTCTTCAATGAATCATC 180  
 DB 121 GAATTAATAATGGAATCTTTATCGAATCCAGCTGATTTTGTCTTCTTCAATGAATCATC 180

QY 181 TCTCTAAAGTGGAAATTTTGTAAAGAGAAAGATCTGAAGTCTGTAGAGGCTTAGTGATG 240  
 DB 181 TCTCTAAAGTGGAAATTTTGTAAAGAGAAAGATCTGAAGTCTGTAGAGGCTTAGTGATG 240

QY 241 GAGACAAATTCGTCTGGAGAGATCTCGTTATTAAAGACTCGGAAGCCATATACGATAACA 300  
 DB 241 GAGACAAATTCGTCTGGAGAGATCTCGTTATTAAAGACTCGGAAGCCATATACGATAACA 300

QY 301 AAGCAAGCTCAAGGCTGAGTCTGAGGAGACATTAATAGATTTCATTGAAGCTTTGAGGCTT 360  
 DB 301 AAGCAAGCTCAAGGCTGAGTCTGAGGAGACATTAATAGATTTCATTGAAGCTTTGAGGCTT 360

QY 361 TATGGTAGACATCGCAGAAAGATTGAAGAACATGTAGCAACAAACCTGCTGCCAGATA 420  
 DB 361 TATGGTAGACATCGCAGAAAGATTGAAGAACATGTAGCAACAAACCTGCTGCCAGATA 420

QY 421 AGAAGTCACCTCAGAAATTTTCTCCAGGTAGAGAGAGGCTGAAGCTAAAGGTGTA 480  
 DB 421 AGAAGTCACCTCAGAAATTTTCTCCAGGTAGAGAGAGGCTGAAGCTAAAGGTGTA 480

QY 481 GCTATGGTCAAGGCTAGACATAGCTATTCTCTCCAGGCTTAAGCGTAAACCAAC 540  
 DB 481 GCTATGGTCAAGGCTAGACATAGCTATTCTCTCCAGGCTTAAGCGTAAACCAAC 540

QY 541 AATCCCTTATCTCGAAGACGGGAAGTGAACGATCTCTATGTCAAAAACGGGTGTGAAT 600  
 DB 541 AATCCCTTATCTCGAAGACGGGAAGTGAACGATCTCTATGTCAAAAACGGGTGTGAAT 600

DB 541 AATCCCTTATCTCGAAGACGGGAAGTGAACGATCTCTATGTCAAAAACGGGTGTGAAT 600  
 QY 601 GATGGAAAGAGTCCCTTCGATCAGAAAAAGTGTCCGATCTCGATGGCCCAATGAAGAT 660  
 DB 601 GATGGAAAGAGTCCCTTCGATCAGAAAAAGTGTCCGATCTCGATGGCCCAATGAAGAT 660  
 QY 661 CGACAAACAATCAAGCCTGAAGAGAAAACTCTCGAGGAAGACAACTGTTCAGATTGTTTC 720  
 DB 661 CGACAAACAATCAAGCCTGAAGAGAAAACTCTCGAGGAAGACAACTGTTCAGATTGTTTC 720  
 QY 721 ACTCATCAGTATCTCTCTGTGTCATCTCCATGAATAAAGTTGTATAGACATCAAAAC 780  
 DB 721 ACTCATCAGTATCTCTCTGTGTCATCTCCATGAATAAAGTTGTATAGACATCAAAAC 780  
 QY 781 GCAAGCACTTTCCGCGAGTTCTTCGCTTTCAGGGAAGGGAAGTCAAGATAACAGGTA 840  
 DB 781 GCAAGCACTTTCCGCGAGTTCTTCGCTTTCAGGGAAGGGAAGTCAAGATAACAGGTA 840  
 QY 841 AGAAGGAGTCAAACTCAGATTTGAATGCAAAAATCTCTGGAACCGGTAAATGAGCAAGGA 900  
 DB 841 AGAAGGAGTCAAACTCAGATTTGAATGCAAAAATCTCTGGAACCGGTAAATGAGCAAGGA 900  
 QY 901 CCTCAGACTTATCGGATGCATATCCCTGTGTAGTGCATTTGGGGAGCTCAATAAAGT 960  
 DB 901 CCTCAGACTTATCGGATGCATATCCCTGTGTAGTGCATTTGGGGAGCTCAATAAAGT 960  
 QY 961 TCTCTATCAGTCTCTCTTCAGAGCCAGATAGTATCCACACAGTGTGAGGAGATTAT 1020  
 DB 961 TCTCTATCAGTCTCTCTTCAGAGCCAGATAGTATCCACACAGTGTGAGGAGATTAT 1020  
 QY 1021 CAGTCTGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTACTGCC 1080  
 DB 1021 CAGTCTGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTACTGCC 1080  
 QY 1081 GCAACTTTCGCCCTCATCATTTTGGCCTCCCGATTTCTAGTGTGCTCACTGTTCAGGG 1140  
 DB 1081 GCAACTTTCGCCCTCATCATTTTGGCCTCCCGATTTCTAGTGTGCTCACTGTTCAGGG 1140  
 QY 1141 AACTCAGCTCCGAAATCTGGCTGCATCGGCGAGCCACTGTTCAGCTGTGCTAGTCTGG 1200  
 DB 1141 AACTCAGCTCCGAAATCTGGCTGCATCGGCGAGCCACTGTTCAGCTGTGCTAGTCTGG 1200  
 QY 1201 TGGGCTGCCAATGGATTATTACCTTTATGTGCTCTCTTATAGTTTCTAGTGTGCTCACTAGT 1260  
 DB 1201 TGGGCTGCCAATGGATTATTACCTTTATGTGCTCTCTTATAGTTTCTAGTGTGCTCACTAGT 1260  
 QY 1261 CATCTCTCATCTACTTTTGGACCATCATGTGTGATGTAGAGTACACAAAGCAAGCACTTTA 1320  
 DB 1261 CATCTCTCATCTACTTTTGGACCATCATGTGTGATGTAGAGTACACAAAGCAAGCACTTTA 1320  
 QY 1321 CAACATGGTTCTGTGACAGCCGAGAGCAAGACACTCCGAGGCTCAAAAGGCTCGATCT 1380  
 DB 1321 CAACATGGTTCTGTGACAGCCGAGAGCAAGACACTCCGAGGCTCAAAAGGCTCGATCT 1380  
 QY 1381 TCATCGACTCAGAGGATTTGAAAAATAAGATTAACAGTAAACAGTTTGTCTATGAGCAGCTTCT 1440  
 DB 1381 TCATCGACTCAGAGGATTTGAAAAATAAGATTAACAGTAAACAGTTTGTCTATGAGCAGCTTCT 1440  
 QY 1441 GCAACACCTCAGAGTGTGCAAAAGGTTTTCAGATGGAGCAGAGACAGAAACAAAGTTGAC 1500  
 DB 1441 GCAACACCTCAGAGTGTGCAAAAGGTTTTCAGATGGAGCAGAGACAGAAACAAAGTTGAC 1500  
 QY 1501 CGGTCTCTGTGTGGCTCAAAACACTCCGTCCAGTAGTGTATGTGTAGGCGGATGCATCA 1560  
 DB 1501 CGGTCTCTGTGTGGCTCAAAACACTCCGTCCAGTAGTGTATGTGTAGGCGGATGCATCA 1560  
 QY 1561 GAAAGCAAGAGATGGCCAAATGGTGGAGTGAAGAAAAAGAAAGCACTTAATAAAA 1620  
 DB 1561 GAAAGCAAGAGATGGCCAAATGGTGGAGTGAAGAAAAAGAAAGCACTTAATAAAA 1620  
 QY 1621 CCTCAAACTTCAGAGTCCCAATGACCGCGCAGTAGAATCAGCTTCAATATTAACCGATCCA 1680  
 DB 1621 CCTCAAACTTCAGAGTCCCAATGACCGCGCAGTAGAATCAGCTTCAATATTAACCGATCCA 1680









Dh 1261 CATCTCCATCTACTTTTGGACCATCATGTGATGTAGAGTACACAAAGCAAGCACTTTA 1320  
Qy 1321 CAACATGGTTCTGTGTCAGAGCCGAGAGCAAGAACACATCCGAGGCATCAAGGCTCGATCT 1380  
Db 1321 CAACATGGTTCTGTGTCAGAGCCGAGAGCAAGAACACATCCGAGGCATCAAGGCTCGATCT 1380  
Qy 1381 TCACCTGGACTCAGAGGATGTTGAAAATAAGAGTAACACAGTTTGTGTCATGAGCAGCTTCT 1440  
Db 1381 TCACCTGGACTCAGAGGATGTTGAAAATAAGAGTAACACAGTTTGTGTCATGAGCAGCTTCT 1440  
Qy 1441 GCAACACCTGAGAGTGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAACAAGTTGCAC 1500  
Db 1441 GCAACACCTGAGAGTGATGCAAAAGGTTTCAGATGGAGCAGAGACAGAAAACAAGTTGCAC 1500  
Qy 1501 CGGTCTCTGTCGCTCAAAACACTCCGTCGAGTAGTGATGTTGAGCGCGATGCATCA 1560  
Db 1501 CGGTCTCTGTCGCTCAAAACACTCCGTCGAGTAGTGATGTTGAGCGCGATGCATCA 1560  
Qy 1561 GAAAGGCAAGAGGATGGCAACCAATGGTGAGGTGAAGAAACGAATGAAGACACTAATAAA 1620  
Db 1561 GAAAGGCAAGAGGATGGCAACCAATGGTGAGGTGAAGAAACGAATGAAGACACTAATAAA 1620  
Qy 1621 CCTCAAACTTCAGAGTCCAAATGCACCGCCGAGTAGAATCAGCTCCCAATATAACCGATCCA 1680  
Db 1621 CCTCAAACTTCAGAGTCCAAATGCACCGCCGAGTAGAATCAGCTCCCAATATAACCGATCCA 1680  
Qy 1681 TGGAGTCTGTGTCGACGAGGTCGAATTCGCTTCCAAAGTCTCTTCTCCAGAGAGGTA 1740  
Db 1681 TGGAGTCTGTGTCGACGAGGTCGAATTCGCTTCCAAAGTCTCTTCTCCAGAGAGGTA 1740  
Qy 1741 TTGCGCGAAAGTTTACATATCGAGAGAACACACAGAGGAGGAACACACACACACAGAA 1800  
Db 1741 TTGCGCGAAAGTTTACATATCGAGAGAACACACAGAGGAGGAACACACACACACAGAA 1800  
Qy 1801 CAAAGATATCCAAATGGCACTTGATCTTAACTTTCACAGCTCAGTTAAACACCGATGAT 1860  
Db 1801 CAAAGATATCCAAATGGCACTTGATCTTAACTTTCACAGCTCAGTTAAACACCGATGAT 1860  
Qy 1861 CAAGAGAGAGAGAGAACACAGGATTTCTTGGAAATCGGATTCGATGCTTCAAGCTAATG 1920  
Db 1861 CAAGAGAGAGAGAGAACACAGGATTTCTTGGAAATCGGATTCGATGCTTCAAGCTAATG 1920  
Qy 1921 AGTAGAGAGAACACAGGTTTAAACCATACAAAGATGTTCCATGGAGCCAAAGAAAGT 1980  
Db 1921 AGTAGAGAGAACACAGGTTTAAACCATACAAAGATGTTCCATGGAGCCAAAGAAAGT 1980  
Qy 1981 AGAATCTCTCAACAAATCTCTATCATTCATGTGGAACAGAAAGATCCCAACCGATGCGG 2040  
Db 1981 AGAATCTCTCAACAAATCTCTATCATTCATGTGGAACAGAAAGATCCCAACCGATGCGG 2040  
Qy 2041 TTGGAAGTCAAGCTTCCAGATGAGACTCTATTTTTCATCTGATCTGTTGTTGTTGTTGTTG 2100  
Db 2041 TTGGAAGTCAAGCTTCCAGATGAGACTCTATTTTTCATCTGATCTGTTGTTGTTGTTGTTG 2100  
Qy 2101 TTTTAAAGTTTTCAGACCACTGCTACATTTTCTTTTCTTTTTCAGGCTTGTGTTGTTGTTGTTG 2160  
Db 2101 TTTTAAAGTTTTCAGACCACTGCTACATTTTCTTTTCTTTTTCAGGCTTGTGTTGTTGTTGTTG 2160  
Qy 2161 TTCCTGTGTCATGCTTCTCTGTAACATTTGACTCTGTTATTTCAACAATATCAAACT 2220  
Db 2161 TTCCTGTGTCATGCTTCTCTGTAACATTTGACTCTGTTATTTCAACAATATCAAACT 2220  
Qy 2221 GTTTAAATCTTTTTTTTCCA 2240  
Db 2221 GTTTAAATCTTTTTTTTCCA 2240

RESULT 4  
AAD05773  
ID AAD05773 standard; cDNA; 2240 BP.  
XX  
AC AAD05773;  
XX

DT 31-JUL-2001 (first entry)  
XX Arabidopsis thaliana transcription factor, G214 cDNA.  
DE Transcription factor; biochemical characteristic; controlling element;  
XX structural characteristic; developmental characteristic; gene therapy;  
KW agricultural biotechnology; plant trait modification; ss.  
KW Arabidopsis thaliana.  
XX OS Arabidopsis thaliana.  
XX FH Key Location/Qualifiers  
FT CDS 238..2064  
FT /\*tag= a  
FT /product= "Transcription factor, G214"  
XX  
XX W0200136597-A1.  
XX 25-MAY-2001.  
XX 14-NOV-2000; 2000WO-US31344.  
XX 17-NOV-1999; 99US-0166228.  
PR 17-APR-2000; 2000US-0197899.  
PR 22-AUG-2000; 2000US-0227439.  
XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
PA (CREE/) CREELMAN R.  
PA (YUGG/) YU G.  
PA (ADAM/) ADAM L.  
PA (RIEC/) RIECHMANN J L.  
PA (HEAR/) HEARD J.  
PA (SAMA/) SAMAHA R.  
PA (PILG/) PILGRIM M.  
PA (PINE/) PINEDA O.  
PA (JIAN/) JIANG C.  
XX  
XX Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;  
PI Pilgrim M, Pineda O, Jiang C;  
XX  
XX WPI; 2001-335999/35.  
DR P-PSDB; AAE01889.  
XX  
XX Nucleic acids encoding plant transcription factor polypeptides, useful  
PT for altering the biochemical characteristics of plants e.g. corn,  
PT potato and cotton plants -  
XX  
XX Claim 4; Page 55-57; 127pp: English.  
XX  
XX The present sequence is Arabidopsis thaliana transcription factor,  
CC G214 cDNA. The transcription factor is used for altering a plant's  
CC biochemical characteristics. The transcription factor may be used to  
CC alter the structure and developmental characteristics of plants such as  
CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,  
CC alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry,  
CC raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant,  
CC grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,  
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,  
CC rosaceous fruits and/or vegetable brassicas. Transcription factors are  
CC key controlling elements of biological pathways and altering expression  
CC levels of 1 or more transcription factors can change entire biological  
CC pathways in an organism. Therefore manipulating transcription factor  
CC levels in plants offers great potential in agricultural biotechnology  
CC for modifying a plant's traits. Transcription factor cDNA is useful in  
CC gene therapy.  
XX  
SQ Sequence 2240 BP; 685 A; 473 C; 477 G; 605 T; 0 other;

Query Match 99.4%; Score 2240; DB 22; Length 2240;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TGAGATTTCCTCCATTCCTGAGTCTCTGTTCTCTTTCTTTGTTCTATGATCAAAAGCA 60  
|||||



```

Qy 2221 GTTTAAATCTTTTTCCTCA 2240
Db 2221 GTTTAAATCTTTTTCCTCA 2240

RESULT 5
ABZ13695
ID ABZ13695 standard; DNA; 1827 BP.
XX AC ABZ13695;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1500.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN W0200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US26685.
XX PR 24-AUG-2000; 2000US-227866P.
XX PR 26-JAN-2001; 2001US-264647P.
XX PR 22-JUN-2001; 2001US-300111P.
XX PA (SCRI ) SCRIPPS RES INST.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Harper JF, Kreps J, Wang X, Zhu T;
XX PI WPI; 2002-304127/34.
XX DR
XX PT Identifying a stress condition to which a plant cell has been exposed
XX PT and producing plants with increased tolerance to these abiotic stresses
XX PT
XX PS Claim 144; SEQ ID NO 1500; 577pp + Sequence Listing; English.
XX PS
XX CC The invention relates to identifying a stress condition to which a plant
XX CC cell has been exposed, comprising:
XX CC (a) contacting nucleic acid representative of expressed polynucleotides
XX CC in the plant cell with an array or probes representative of the plant
XX CC cell genome; and
XX CC (b) detecting a profile of expressed polynucleotides in the plant cell
XX CC characteristic of a stress response. The method is useful in the
XX CC production of transgenic plants, cells and seeds and in producing plants
XX CC with increased tolerance to abiotic stress. The present sequence is that
XX CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX CC in methods of the invention.
XX CC Note: The sequence data for this patent is not represented in the printed
XX CC specification but is based on sequence information supplied to Derwent by
XX CC the European Patent Office.
XX SQ Sequence 1827 BP; 591 A; 397 C; 419 G; 420 T; 0 other;

Query Match 81.1%; Score 1827; DB 24; Length 1827;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 ATGAGACAAATTCGTGGAGAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA 297
Db 1 ATGAGACAAATTCGTGGAGAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA 60

Qy 298 ACAAGCAAGCTCAAGAGTGGACTGAGGAAGACATAATAGATTCAATTGAAGCTTTGAGG 357
Db 61 ACAAGCAAGCTCAAGAGTGGACTGAGGAAGACATAATAGATTCAATTGAAGCTTTGAGG 120

Qy 358 CTTTATGGTAGACATGGCAGAGAGATTGAAGAACATGTACCAACAAACCTGCTGCCAG 417
Db CTTTATGGTAGACATGGCAGAGAGATTGAAGAACATGTACCAACAAACCTGCTGCCAG 1260

```

QY 1498 GACCGTCTCGTGTGCTCAACACATCCGTCGAGTAGTGATGATGTTGAGCGGATGCA 1557  
 Db 1261 GACCGTCTCGTGTGCTCAACACATCCGTCGAGTAGTGATGATGTTGAGCGGATGCA 1320  
 QY 1558 TCAGAAAGGCAAGAGGATGCGCAATGGTGAGTGAAGAAACGAATGAAGACACTAAT 1617  
 Db 1321 TCAGAAAGGCAAGAGGATGCGCAATGGTGAGTGAAGAAACGAATGAAGACACTAAT 1380  
 QY 1618 AAACCTCAAACTTCAGAGTCCAATGTCACGCGCAGTAGAATCAGCTCCAATATAACCGAT 1677  
 Db 1381 AAACCTCAAACTTCAGAGTCCAATGTCACGCGCAGTAGAATCAGCTCCAATATAACCGAT 1440  
 QY 1678 CCATGGAAGTCTGTCTGACGAGGTCGAATTCCTTCCAAGCTCTCTTCCAGAGAG 1737  
 Db 1441 CCATGGAAGTCTGTCTGACGAGGTCGAATTCCTTCCAAGCTCTCTTCCAGAGAG 1500  
 QY 1738 GTATTGCGCAAAAGTTTTACATATCGAGAAGAACACAGAGAGGAAGAACACACACAA 1797  
 Db 1501 GTATTGCGCAAAAGTTTTACATATCGAGAAGAACACAGAGAGGAAGAACACACAA 1560  
 QY 1798 GAACAAAGATATCCAATGGCACTTGATCTTTAACTTCACAGCTCAGTTAACACCACTTGAT 1857  
 Db 1561 GAACAAAGATATCCAATGGCACTTGATCTTTAACTTCACAGCTCAGTTAACACCACTTGAT 1620  
 QY 1858 GATCAAGAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATAGATGCTTCAAAGCTA 1917  
 Db 1621 GATCAAGAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATAGATGCTTCAAAGCTA 1680  
 QY 1918 ATCAGTAGAGAGAGAGAGAGGTTTTAAACCATACAGAAAGATGTTCCATGGAAGCCAAAGAA 1977  
 Db 1681 ATCAGTAGAGAGAGAGAGAGGTTTTAAACCATACAGAAAGATGTTCCATGGAAGCCAAAGAA 1740  
 QY 1978 AGTAGAATCCTCAACCAATCTCTATCATTCATGTGGAACAGAGAAATCCCAACCGGATG 2037  
 Db 1741 AGTAGAATCCTCAACCAATCTCTATCATTCATGTGGAACAGAGAAATCCCAACCGGATG 1800  
 QY 2038 CGGTTGGAAGTCAAGCTCCACATGA 2064  
 Db 1801 CGGTTGGAAGTCAAGCTCCACATGA 1827

RESULT 6  
 AAV65381  
 ID AAV65381 standard; DNA; 4344 BP.  
 AC  
 XX  
 AC AAV65381;  
 XX  
 DT 15-FEB-1999 (first entry)  
 XX  
 DE Arabidopsis phytochrome regulated transcription factor CCA1 DNA.  
 XX  
 KW Phytochrome regulated transcription factor; CCA1;  
 KW chlorophyll binding protein; Lhcb1\*3; circadian rhythm; flowering;  
 KW transgenic plant; ss.  
 XX  
 OS Arabidopsis thaliana ecotype Columbia.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1332..4043  
 FT FT /\*tag= a  
 FT FT /note= "contains introns"  
 FT FT 1..1999  
 FT FT /\*tag= b  
 FT FT /number= 1  
 FT FT 1200..1283  
 FT FT /\*tag= c  
 FT FT /number= 1  
 FT FT 1332..1370  
 FT FT /\*tag= d  
 FT FT /number= 2  
 FT FT 1371..1448  
 FT FT /\*tag= e

FT exon /number= 2  
 1449..1560  
 /\*tag= f  
 FT intron /number= 3  
 1561..1648  
 /\*tag= g  
 FT exon /number= 3  
 1649..1710  
 /\*tag= h  
 FT intron /number= 4  
 1711..2189  
 /\*tag= i  
 FT exon /number= 4  
 2190..2384  
 /\*tag= j  
 FT intron /number= 5  
 2385..2471  
 /\*tag= k  
 FT exon /number= 5  
 2472..2645  
 /\*tag= l  
 FT intron /number= 6  
 2646..2728  
 /\*tag= m  
 FT exon /number= 6  
 2729..3610  
 /\*tag= n  
 FT intron /number= 7  
 3611..3680  
 /\*tag= o  
 FT exon /number= 7  
 3681..4043  
 /\*tag= p  
 /\*number= 8  
 WO9848007-A1.  
 XX  
 PD 29-OCT-1998.  
 XX  
 PF 17-APR-1998; 98WO-US07693.  
 XX  
 PR 18-APR-1997; 97US-0843572.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 DR WPI; 1998-583656/49.  
 DR P-PSDB; AAW79280.  
 XX  
 PT Alteration of plant circadian rhythms - by transformation with a  
 PT transcription factor for the promoter of the Arabidopsis chlorophyll  
 PT binding protein  
 XX  
 PS Claim 1; Page 36-40; 47pp; English.  
 XX  
 CC This genomic DNA sequence codes for a phytochrome-regulated  
 CC transcription factor, designated CCA1 (see AAW79280), that binds to  
 CC the promoter region of the chlorophyll binding protein gene  
 CC (Lhcb1\*3) of Arabidopsis thaliana. The genomic clone was isolated  
 CC by screening an Arabidopsis ecotype Columbia genomic library using  
 CC CCA1 cDNA (see AAV65382) as probe. A claimed method for altering  
 CC plant response to daylength comprises transforming a plant with  
 CC either the CCA1 genomic DNA or cDNA, a hybridising sequence or a  
 CC nucleic acid encoding a protein containing amino acids 24-75 of the  
 CC 608-amino acid CCA1 protein. Manipulation of plant circadian  
 CC rhythms (e.g. by overexpression of CCA1) can be used to delay  
 CC flowering, and thus increase yield from both forage crops, where  
 CC biomass will be produced for longer, and from seed crops such as  
 CC rape, where a longer time in vegetative state will lead to larger  
 CC plants and thus a higher seed set (claimed).  
 XX  
 SQ Sequence 4344 BP; 1303 A; 808 C; 883 G; 1350 T; 0 other;  
 Query Match 67.8%; Score 1527.6; DB 19; Length 4344;

Best Local Similarity 87.9%; Pred. No. 0;			
Matches 1803; Conservative 0; Mismatches 9; Indels 240; Gaps 3;			
Qy	441	TTTTCTCAAGGTAGAGAAAGGCTGAAGCTAAAGGTAGCTATGGGTCAAGCGCTAGA	500
Db	2180	TCTCTCAGGTAGAGAAAGGCTGAAGCTAAAGGTAGCTATGGGTCAAGCGCTAGA	2239
Qy	501	CATAGCTATTCTCTCTCACGGCCTTAGCGTAAACCAAACTCTTATCTCGAAAGAC	560
Db	2240	CATAGCTATTCTCTCTCACGGCCTTAGCGTAAACCAAACTCTTATCTCGAAAGAC	2299
Qy	561	GGGAAGTGGAAACGATCTTATGTCAAAAACGGGTGTAATCATGGAAGAGTCCCTGG	620
Db	2300	GGGAAGTGGAAACGATCTTATGTCAAAAACGGGTGTAATCATGGAAGAGTCCCTGG	2359
Qy	621	ATCAGAAAAAGTGTGCAATCTG-----	643
Db	2360	ATCAGAAAAAGTGTGCAATCTGAGGTGATTTTCATGGTCATATGGCATCTTTTTCAGT	2419
Qy	644	-----AGATGGCCAA	653
Db	2420	GTGTCAANTGCTCTCATGTTATTAATACAGATTGTGTGCTTCGTTTATAGATGGCAA	2479
Qy	654	TGAAGATCGCAACAACTCAAAAGCCTGAAGAGAAAACTCTGCAGGAAGACAACTGTTGAGA	713
Db	2480	TGAAGATCGCAACAACTCAAAAGCCTGAAGAGAAAACTCTGCAGGAAGACAACTGTTGAGA	2539
Qy	714	TTGTTTCACTCATCAGTATCTCTGTGTCATCTCCATGAATAAAAGTTGTATAGAGAC	773
Db	2540	TTGTTTCACTCATCAGTATCTCTGTGTCATCTCCATGAATAAAAGTTGTATAGAGAC	2599
Qy	774	ATCAACCGCAAGCACTTCCGCGAGTCTTCCCTTTCAGCGGAAG-----	817
Db	2600	ATCAACCGCAAGCACTTCCGCGAGTCTTCCCTTTCAGCGGAAGGTTAAACAACTCT	2659
Qy	818	-----	817
Db	2660	TTCAATGCTATTGAGGTTTTAAGACGATTAGTACTTTTCATGAACCTAAACCGTGGGG	2719
Qy	818	-----AGGAAGTCAGATAACAGGGTGAAGAAAGAGTCAAACTCAGATTTGAATGCA	870
Db	2720	GAATAACAGGGAAGTCAAGATAACAGGGTGAAGAAAGAGTCAAACTCAGATTTGAATGCA	2779
Qy	871	AAATCTCTGAAACCGTAAATGACCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG	930
Db	2780	AAATCTCTGAAACCGTAAATGACCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG	2839
Qy	931	CTAGTGCATTGGGAGCTCAATAACAAAGTTCTCTATCAATCCTCCTTCAGAGCCAGAT	990
Db	2840	CTAGTGCATTGGGAGCTCAATAACAAAGTTCTCTATCAATCCTCCTTCAGAGCCAGAT	2899
Qy	991	AGTCATCCCAACAGATTGAGGAGATTATCAGTCGTTTCTTAATCATATAATGTCAACC	1050
Db	2900	AGTCATCCCAACAGATTGAGGAGATTATCAGTCGTTTCTTAATCATATAATGTCAACC	2959
Qy	1051	CTTTTCAAAACACGGCTCTTTTATCTACGCGCAACTTTCCGCTCATATTTGGCCTCCC	1110
Db	2960	CTTTTCAAAACACGGCTCTTTTATCTACGCGCAACTTTCCGCTCATATTTGGCCTCCC	3019
Qy	1111	GATTCTAGTGGTGGCTCACCTGTTTCCAGGGAACCTCACCTCGAATCTGGCTGCCATGGCC	1170
Db	3020	GATTCTAGTGGTGGCTCACCTGTTTCCAGGGAACCTCACCTCGAATCTGGCTGCCATGGCC	3079
Qy	1171	GCAGCCACTGTTGCAAGTGTAGTGTGTTGGTGGCTGCCAATGATTTACCTTTATGT	1230
Db	3080	GCAGCCACTGTTGCAAGTGTAGTGTGTTGGTGGCTGCCAATGATTTACCTTTATGT	3139
Qy	1231	GCTCCTCTTAGTTTCAAGTGTGTTTCACTAGTCATCTCCATCTACTTTGGACCATCATGT	1290
Db	3140	GCTCCTCTTAGTTTCAAGTGTGTTTCACTAGTCATCTCCATCTACTTTGGACCATCATGT	3199
Qy	1291	GATGTAGAGTACAAAAAGCAAGCACTTTTCAACATGGTTCTGTGAGAGCCGAGAGCAA	1350

Db	3200	GATGTAGAGTACAAAAAGCAAGCACCTTTACAACATGGTTCTGTGCAGAGCCGAGACAA	3259
Qy	1351	GAACACTCCGAGGATCAAAAGCTCGATCTTCTACTTGGACTCAGAGGATGTTGAAATTAAG	1410
Db	3260	GAACACTCCGAGGATCAAAAGCTCGATCTTCTACTTGGACTCAGAGGATGTTGAAATTAAG	3319
Qy	1411	AGTAAACAGTTTGTCTAGTACGACCTTCTCAACACCTCAGAGTGTGATGCAAGGTTCA	1470
Db	3320	AGTAAACAGTTTGTCTAGTACGACCTTCTCAACACCTCAGAGTGTGATGCAAGGTTCA	3379
Qy	1471	GATCGAGCAGGACAGAAAACAAAGTTGACCGGTCTCTGTGGCTCAAACTCCGTCG	1530
Db	3380	GATCGAGCAGGACAGAAAACAAAGTTGACCGGTCTCTGTGGCTCAAACTCCGTCG	3439
Qy	1531	AGTAGTGTATGTGTGAGGCGCATGCAATAAGGCAAGAGGATGGCAACCAATGGTGAG	1590
Db	3440	AGTAGTGTATGTGTGAGGCGCATGCAATAAGGCAAGAGGATGGCAACCAATGGTGAG	3499
Qy	1591	GTGAAAGAAACGAATGAAGACATAATAAACTCAAACTTCAAGTGTCAATGCACCGCG	1650
Db	3500	GTGAAAGAAACGAATGAAGACATAATAAACTTCAAACTTCAAGTGTCAATGCACCGCG	3559
Qy	1651	AGTAGAATCAGCTCCATATATAACCGATCCATGGAAGTCTGTCTGACGA-----	1700
Db	3560	AGTAGAATCAGCTCCATATATAACCGATCCATGGAAGTCTGTCTGACGA-----	3619
Qy	1701	-----	1700
Db	3620	TGGACTAAAGATCAACTTCTTATTTCAAAATCATTTTCTCATATAAATATTGTACATTC	3679
Qy	1701	GGGTGGAATTTGCTTCCAAAGCTCTTCTCCAGAGAGTATTTGCCGCAAGTTTTCACATA	1760
Db	3680	GGGTGGAATTTGCTTCCAAAGCTCTTCTCCAGAGAGTATTTGCCGCAAGTTTTCACATA	3739
Qy	1761	TCGAGAAGAACACAGAGAGGAGAAACAACAACAAGAGATATCCAAATGGCACT	1820
Db	3740	TCGAGAAGAACACAGAGAGGAGAAACAACAACAAGAGATATCCAAATGGCACT	3799
Qy	1821	TGATCTTTAACTTTCACAGCTCAGTTTAAACAACAGTTTGTATCAAGAGGAGAGAAACAC	1880
Db	3800	TGATCTTTAACTTTCACAGCTCAGTTTAAACAACAGTTTGTATCAAGAGGAGAGAAACAC	3859
Qy	1881	AGGATTTCTTGGAAATCGGATTTAGATGCTTCAAGCTAAATAGTAGAGGAGAAACAGGTTT	1940
Db	3860	AGGATTTCTTGGAAATCGGATTTAGATGCTTCAAGCTAAATAGTAGAGGAGAAACAGGTTT	3919
Qy	1941	TAAACCATACAAAGAGTGTTCCTTCCATGGAAGCCAAAGAAAGTGAATCTCTCAACAATCC	2000
Db	3920	TAAACCATACAAAGAGTGTTCCTTCCATGGAAGCCAAAGAAAGTGAATCTCTCAACAATCC	3979
Qy	2001	TATCATTTCATGTGGAAACAGAAAGATCCCAACCGGATCGGTTGGAAACTCAAGCTTCCAC	2060
Db	3980	TATCATTTCATGTGGAAACAGAAAGATCCCAACCGGATCGGTTGGAAACTCAAGCTTCCAC	4039
Qy	2061	ATGAGACTCTATTTTCACTGTATCTGTGTTGTGTTGTGTTGTGTTGTGTTGTGTTGTGTT	2120
Db	4040	ATGAGACTCTATTTTCACTGTATCTGTGTTGTGTTGTGTTGTGTTGTGTTGTGTTGTGTT	4099
Qy	2121	CTGCTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	2180
Db	4100	CTGCTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	4159
Qy	2181	TGTAACATTTGACTCTGTATTTTCAACAAATCATAACTGTTTAACTTTTTTTTTTCCCA	2240
Db	4160	TGTAACATTTGACTCTGTATTTTCAACAAATCATAACTGTTTAACTTTTTTTTTTCCCA	4219
Qy	2241	AAAAAANAANA 2252	
Db	4220	ACCTGGAAGAA 4231	

ID AAV09296 standard; DNA; 2526 BP.

XX AC AAV09296;

XX DT 07-JUL-1998 (first entry)

XX DE Nucleotide sequence of LHY gene.

XX KW Late elongated hypocotyl gene; LHY gene; flowering; day length; transgenic plant; promoter; circadian oscillatory; ss.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers  
CDS 337..2275

XX FT /\*tag= a  
FT /product= "LHY protein"

XX PN WO9749811-A1.

XX PD 31-DEC-1997.

XX PF 23-JUN-1997; 97WO-GB01676.

XX PR 21-JUN-1996; 96GB-0013132.

XX PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.

XX PI Coupland GW, Schaffer RJ;

XX DR WPI; 1998-077174/07.

XX DR P-PSDB; AAW42084.

XX PT New isolated late elongated hypocotyl gene - used to obtain plants with altered flowering characteristics, and for expression of sequences under circadian rhythm control

XX PS Claim 4; Fig 1; 90pp; English.

XX CC This nucleotide sequence is of the late elongated hypocotyl (LHY) gene, which is one of the genes that controls the timing of flowering in relation to the length of day. The LHY type genes and derivatives can be used for producing transgenic plants in which a physical characteristic, such as a flowering characteristic which may include timing of flowering, is affected. The LHY promoters can be used to regulate expression of a sequence of interest in a circadian oscillatory manner.

XX SQ Sequence 2526 BP; 767 A; 523 C; 528 G; 708 T; 0 other;

Query Match 9.6%; Score 217.2; DB 19; Length 2526;  
Best Local Similarity 49.9%; Pred. No. 6.3e-46;  
Matches 986; Conservative 0; Mismatches 823; Indels 165; Gaps 11;

QY 235 GTGATGGAGACAAATTCGTCTGGAAGATCTGGTTATTAGACTCGGAAGCCATATACG 294  
DB 335 GTTATGGATACTAATACATCTGGAGAAGAAATTATTAGCTAAGCAAGAAAGCCATATACA 394  
QY 295 ATAACAAAGCAACGTTGAAAGGTGGACTGAGGAAGAACATAATAGATTTCATTGAAGCTTTG 354  
DB 395 ATAACAAAGCAGCAGAGCGATGGACTGAGGATGACCATGAGAGTTCTAGAACCTTG 454  
QY 355 AGGCTTTTATGGTAGACATGGCAGAAAGATTGAAGAACATGTAGCAACAAAACTGCTGTC 414  
DB 455 AGGCTTTTATGGAAGAGCTTGGCAACGAATTGAAGAACATATTGGGACAAAGACTGCTGTT 514  
QY 415 CAGATAAGAAGTCAACGCTCAGAAATTTTCTCCAAAGGTAGAGAAAGGCTGAAGCTAAA 474  
DB 515 CAGATCAGAAGTCATGCACAAAGTCTTCACAAAGTTGGAGAAAGAGGCTGAAGTTAAA 574  
QY 475 GGTGTAGCTATGGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCTTAAGCGTAAA 534  
DB 575 GGCATCCCTGTTTGGCAAGCTTTGGACATAGAAATTCGGCTCTCTGCTCTTAACCAAAA 634

QY 535 CCAAAACAATCCTTATCTCTCGAAGACGGGAAGTGAACGATCCTTATGTCAAAAACGGGT 594  
DB 635 CCCAATACTCCTTATCTCTCGAAGACCTGGGAACAACGGTACATCTTCTCTCAAGATCA 694  
QY 595 GTGAATGATGGAAGAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCTGAGATGGCCAAT 654  
DB 695 TCAGCAAAAAGATGCAAAAACCTTGTTCATCGGCTCTTCTTACAGATTGAATCAGGCGTTC 754  
QY 655 GAAGATCGACAACAATCAAAAGCCTGAAGAGAAAAC-----TCTGCAGGAA 699  
DB 755 TTGGATTTGGAANAATGCGGTTCTCTGAGAAAACATCAACTCGAANAAGAAAATCAAGAT 814  
QY 700 GAAACTGTTCAGATTGTTTCACTCATCATGATATCTCTCTGCTGCATCTCCATGAATAAAA 759  
DB 815 GAGAAATGCTCGGTGTTTCTACTGTGAACAAGTATCCCTTTACCAACGAACAGGTAAGT 874  
QY 760 AGTTGTATAGACATCAAAACGCAAGCATT-----CCGGAGTTCTTGGCTTC 809  
DB 875 GCGCATTTGAAACAAGTAAGACCTCAACTGTGGACAACGCGGTTTCAAGATGTTCCCAAG 934  
QY 810 ACGGGAAGAGGGAAGTCAGAATAACAGGGTA-----AGAAAGGAGTCAAAAC 855  
DB 935 AAGAACAAGACAAAGATGGTAACGATGGTACTACTGTGCACAGCATGCAAAACTACCCT 994  
QY 856 TCAGATTGAATGCAAAATCTCTGGAACAACGGTAATGAGCAAGGACCTCAGACTTTATCC- 914  
DB 995 TGGCATTTCCACGACAGATATTGTGAACGGGAATATAGCAAAATGCCCTCAAAATCATCCC 1054  
QY 915 -----G 915  
DB 1055 TCAGGTATGGTATCTCAAGACTTTCATGTTTTCATCTCTATGAGAGAAGAACTCAGGCGAC 1114  
QY 916 ATGCATATCCCTGTGCTAGTGCCATTGGGGAGCTCAATAACAAGTTCTCTATCACAATCCT 975  
DB 1115 GCAATCTTCAAGCTACAACAGCATCTGCTACTACTACAGCTTCTCATCAAGCGTTTCCA 1174  
QY 976 CTTTCAGAGCCAGATAGTATCATCCCAACACAGTTCGAGGAGATTTATCAGTCGTTTCTTAAT 1035  
DB 1175 GCTTGTGCTTCAAGATGATTACCGTTGTTTCTCCAGATATCATCTACTTCTCCAAT 1234  
QY 1036 CATATAATGTCACCCCTTTTACAACACCGGCTCTTTATACCTGCGCAACTTTTGGCTCA 1095  
DB 1235 CTTATTATGTCAACTCTCTCAGAAATCTCGAGCTCATGCTGCAGCTACATTGCGTCT 1294  
QY 1096 TCATTTTGGCCTCCCG-----ATTCTAGTGGTGGCTCACCTGTTTCAGGGAAC 1143  
DB 1295 TCGGTCTGGCCTTATGCGAGTGTGGGAATTTCTGGTGATTTCATCAACCCCAATGAGCTCT 1354  
QY 1144 TCACCTCCGAATCTGGCTGCCATGCGCCAGCCACTGTTTGCAGCTGCTAGTGTCTTGGTGG 1203  
DB 1355 TCTCTCCAAAGTATAACTGCCATTGCCGCTGCTACAGTAGCTGCTGCAACTGTTGGTGG 1414  
QY 1204 GCTGCAATGGATTATTACCTTTATGCTCTCTTCTTCTAGTTCAGGTGGTTTCACTAGTCAT 1263  
DB 1415 GCTTCTCATGAGCTTCTTCTCTGTATGCGCTCCAGCTCCAAATAACATGTTTCTTCA 1474  
QY 1264 CTTCCATCTACTTTTGGACCATCATGTGATGTAGATACACAAAACGACGACTTTACAA 1323  
DB 1475 ACTGTTGAGTTTCAACTCCAGCAATGACTGAAATGGATACCGTTGAAAATACCTCAACCG 1534  
QY 1324 CATGTTTCTGTGACAGCGCAGAGCAAGAACTCCGAGGCGATCAAAAGGCTCGATCTTCA 1383  
DB 1535 TTTGAGAAACAAAACACAGCTCTGCAAGATCAAACTTGGCTTCGAAATCTCCAGTTTCA 1594  
QY 1384 CTGAGCTCAGAGGATGTTGAAAAT-----AAGAGTAAACCAAGTTTGTCTAGAGCAG 1434  
DB 1595 TCATCTGATGATTCAGATGAGACTGGAGTAACCAAGCTAAATGCCGACTCAAAAACCAAT 1654  
QY 1435 CTTCTGCAACACCTGAGAGTGATGCAA-----AGGTTTCAGATGGAGCAGA 1482  
DB 1655 GATGATAAAATTGAGGAGGTTGTTGTTTACTGCGCTGTGTCATGACTCAAAACACTGCCAG 1714



Db 635 CCCAATCTCTTATCTCTGGAACCTGGGAAACCGGTACATCTTCTCTCAAGTATCA 694  
 Qy 595 GTGAATGATGGAAGAGTCCCTTGGATCAGAAAAAGTGTGCAATCTCTGAGATGGCCAAT 654  
 Db 695 TCAGCAAAAGATGCAAAACTGTGTTTCATCGGCTCTTCTTACAGTTGAATCAGCGGTC 754  
 Qy 655 GAAGATCGACAACAACTCAAGCCCTGAAGAGAAAC-----TCTCAGGAA 699  
 Db 755 TTGGAATTTGGAATAATGCGCTTCTCTGAGAAAAACATCAACTGGAAGAAAAATCAAGAT 814  
 Qy 700 GACAATCTGTTGAGATGTTTTCACATCATCAGTATCTCTCTGCTGCATCTCTCATGAATAAA 759  
 Db 815 GAGAAATGCTCGGCTGTTCTACTGTGAACAAGTATCCCTTACCAACGAACACAGTAAAT 874  
 Qy 760 AGTTGTTATAGACATCAAAAGCAACTTT-----CCGAGATTTCTGCTTC 809  
 Db 875 GGCACATTTGAAACAAGTAAGACCTCAACTGTGACCAACGCGTTCAAGATGTTCCCAAG 934  
 Qy 810 ACGGAAGAGGGAAGTCAGATAACAGGTA-----AGNAGGAGTCAAC 855  
 Db 935 AAGAACAAAGCAAGATGTAACGATGGTACTACTGTGCACGATGCAAAACTACCC 994  
 Qy 856 TCAGATTTGAATGCAAAATCTCTGAAAAAGGTAAATGAGCAAGACCTCAGACTTATCC- 914  
 Db 995 TGGCATTTCCACGAGATATTGTGAACGGAATATAGCAAAATGCCCTCAAAATCATCCC 1054  
 Qy 915 -----G 915  
 Db 1055 TCAGGTATGTTATCTCAAGACTTTCATGTTTCATCTATGAGAGAAACCTCACGGCAC 1114  
 Qy 916 ATGCATATCTCTGCTAGTGCCATTCGGGAGCTCAATAACAGTTCTCTATCATCATCT 975  
 Db 1115 GCAAACTTCAAGCTACACAGCAGTCTGCTACTACAGCTTCTCATCAAGCGTTTCCA 1174  
 Qy 976 CTTTCAGAGCCAGATAGTATCTCCCAACACAGTTGCGAGGATATCATGCTGTTCCCTAAT 1035  
 Db 1175 GCTTGTCAATTCAGAGTATACCGTTGCTTCTCCAGATATCATCTACTTCTCCAA 1234  
 Qy 1036 CATATAATGTCAACCTTTTACAAACACCGGCTCTTTATCTGCTGCTAGTCTGCTGG 1203  
 Db 1235 CTTATTATGTCAACTCTCTACAGAACTCTGCAAGTCTGCTGAGCTATGCTGAGCTACATCGCTGT 1294  
 Qy 1096 TCATTTTGGCTCCCG-----ATTCTAGTGGTGGCTCACCTGTTCCAGGAAC 1143  
 Db 1295 TCGGTCTGGCTTATGGAGTGTGGGAATTTCTGTTGATTCATCAACCCCAATGAGCTCT 1354  
 Qy 1144 TCACCTCCGAATCTGGCTGCATCGGCGCAGCCACTGTTGAGCTGCTAGTCTGCTGG 1203  
 Db 1355 TCTCTCCAAAGTATAACTGCCATTCGCGCTGTACAGTAGTCTGCTGCAACTGCTTGGTG 1414  
 Qy 1204 GCTGCCAATGGATATTACCTTTATGCTGCTCTTATGTTTCAAGTGGTTCACCTAGTCAT 1263  
 Db 1415 GCTTCTCATGGACTTCTTCTGTATGCGCTCCAGCTCCAATAACATGTTTCCATTCTCA 1474  
 Qy 1264 CTTCCATCTACTTTTGGACCATCATGTGATGTAGAGTACAAAGCAACACTTTACAA 1323  
 Db 1475 ACTGTTGCACTTCCAATCCAGCAATGACTGAAATGGATACCGTTGAAAAATCTCAACCG 1534  
 Qy 1324 CATGGTTCTGTGAGAGCCGAGAGCAAGAACACTCCGAGGCATCAAAAGGCTCGATCTTCA 1383  
 Db 1535 TTTGAGAAACAAACACAGCTCTGCAAGATCAAACTTGGCTTCGAAATCTCCAGTTCA 1594  
 Qy 1384 CTGGACTCAGAGATGTTGAAAT-----AAGATAAACAGTTTGTCTATGAGCAG 1434  
 Db 1595 TCATCTGTGATGATTCAGATGAGCTGGAGTAACCAAGCTAAATCCGACTCAAAACCAAT 1654  
 Qy 1435 CCTTCTGCAACACTCAGAGTGATGCA-----AGGTTTCAGATGAGCAGGA 1482  
 Db 1655 GATGATAAAATTGAGAGGTTGTTGTTATCGCGCTGTGCTGACTCAAAACACTGCCCCAG 1714  
 Qy 1483 GACAGAAAAAAGTTGACCGGCTCTCTGTTGGCTCAAAACACTCCGTCGAGTAGTGATGAT 1542  
 Db 1715 AAGAAAAATCTTGTGACCGCTCATCTGTGTGGCTCAAAATACACTTTCAGGGAGTGAGCA 1774

Qy 1543 GTTGAGGCGGATGTCATCAGAAAGCAAGAGGATGGCAACATGTGAGGTGAAGAAACG 1602  
 Db 1775 GAAACTGATGCTATTAGATAAAATCGAGAAAGATTAAGAGGATGTCAAGGACACATGAG 1834  
 Qy 1603 AAT-----GAGACACTAATAAACCTTCAAACTTCAGAGTCCAATGCAAGCGCGAGTGA 1656  
 Db 1835 AATCAGCCAGATGTTATTGAGTTAAATAACCGTAAGATTAAATGAGAGACAACACAGC 1894  
 Qy 1657 ATCAGCTCCAATATAAACCGATCCATGGAAGTCTGTCTGACGAGGGTCAATTCGCTTC 1716  
 Db 1895 AACAAACATGCAACTACTGATTCGTGGAAGGAAGTCTCCGAGAGGGTCTGATACGCTTT 1954  
 Qy 1717 CAAGCTCTTCTTCAGAGAGGTTATGCGCAAAAGTTTACATATTCGAGAAGAACACAGA 1776  
 Db 1955 CAGCTCTCTTTGCAAGAGAAAGATTGCTCAAAGCTTTTCCCTCTCAAGTGGCAGAG 2014  
 Qy 1777 GAGGAAGAACAAACAAACAAAGAAAGATATCAATGGCAGCTTGTATTAATCTTCA 1836  
 Db 2015 AATGTGAATAGAAAAACAAAGTGACACGT-----CAATGCCATTTGGCTCTTAATTTCAA 2068  
 Qy 1837 GCTCAGTTAACACCAGTTGATGATCAAGAGGAGAGAGAAACACAGGATTTCTTGAATC 1896  
 Db 2069 AGCCAGGATTTCTGTGCTGCAGACCAAG-----AGGAGTAGTAATGATC 2113  
 Qy 1897 GGATTAGATGCTTCAAAAGCTAATGAGTAGAGAGAAAGACAGGTTTAAACCATACAAAAGA 1956  
 Db 2114 GGTGTTGGAACATGCAAGAGTCTTAAACGAGACAGACAGGATTTAAGCCATACAGAGA 2173  
 Qy 1957 TGTTCATGGAAGCAAGAAAGTAGAATCTCTCAACAACTCTATCATTCATGTGAA 2016  
 Db 2174 TGTTCATGGAAGTGAAGAGAGCAAGTTGGGAACATA-----CAATCAAAGTAT 2227  
 Qy 2017 CAGAAAGATCCCAACCGATCGGTTGGAAGTCAAGCTTCAAGCTTCCACATGAGACTCT 2070  
 Db 2228 GAAAAAGTCTGCAAAAGGCTTCGATTGGAAGGAGAGGCTTCTATCATGACAGACT 2281

## RESULT 9

AAD05761

ID AAD05761 standard; cDNA; 2526 BP.

AC AAD05761;

XX 31-JUL-2001 (first entry)

XX Arabidopsis thaliana transcription factor homologue G680 cDNA.

XX Transcription factor; trait modification; seed characteristic;  
 XX structural characteristic; developmental characteristic; gene therapy;  
 XX agricultural biotechnology; ss.

XX Arabidopsis thaliana.

FH Key Location/Qualifiers

CDS 338..2275

FT /\*tag= a

FT /product= "Transcription factor homologue G680"

XX WO200135727-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000MO-US31457.

XX 17-NOV-1999; 99US-0166228.

XX 17-APR-2000; 2000US-0197899.

XX 22-AUG-2000; 2000US-0227439.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX (REUB/) REUBER L.

XX (CREE/) CREELMAN R.

XX (PILG/) PILGRIM M.









CC to a wild-type or reference plant, or the plant exhibits an altered  
CC phenotype as compared to a wild-type or reference plant, or the plant  
CC exhibits ectopic expression or altered expression of one or more genes  
CC associated with a plant trait as compared to a wild plant. Also included  
CC are a transgenic plant comprising the polynucleotides, a computer  
CC readable medium having stored sequence information, and identifying a  
CC homologue sequence from a database comprising a plurality of known plant  
CC sequences comprising inputting sequence information selected from one of  
CC 464 fully defined sequences given in the specification. The isolated or  
CC recombinant polynucleotide is used for producing a plant having a  
CC modified trait, the method comprising selecting a polynucleotide that  
CC encodes a polypeptide or an antisense nucleic acid, inserting the  
CC polynucleotide or antisense nucleic acid into an expression vector,  
CC introducing the vector into a plant or a cell of a plant to overexpress  
CC the polypeptide or antisense nucleic acid, thereby producing a modified  
CC plant, and selecting for a modified trait (e.g. increased  
CC production of agriculturally useful proteins or metabolic chemicals,  
CC pest tolerance, environmental stress response (e.g. drought), microbial  
CC disease resistance, herbicide resistance, seed and fruit yield, growth  
CC rate, leaf and flower senescence and many other traits listed in the  
CC specification). The present sequence is one of the 232 polynucleotides  
CC encoding an A. thaliana transcription factor.

XX  
SQ Sequence 2526 BP; 767 A; 523 C; 528 G; 708 T; 0 other;

Query Match 9.6%; Score 217.2; DB 24; Length 2526;

Best Local Similarity 49.9%; Pred. No. 6.3e-46;

Matches 986; Conservative 0; Mismatches 823; Indels 165; Gaps 11;

QY	235	GTGATGAGACAAATTCGTCTGGAGAGATCTGGTTATTAGACTCGGAGCCATATACG	294
DB	335	GTATGGATATACTATCATCTGGAGAGAAATTAATAGCTAAGGCAAGAAAGCCATACA	394
QY	295	ATAACAAAGCAACCTGAAAGGTGACTGAGGAAGCAATAATAGATTCATTGAAGCTTG	354
DB	395	ATAACAAAGCAGCAGAGCGATGGACTGAGATGAGCATGAGAGTTCTTAGAAGCCTTG	454
QY	355	AGGCTTTATGGTAGAGCATGCGAAGATTGAAGAAACATGTAGCAACAAACCTGTGTC	414
DB	455	AGGCTTTATGGAAGAGCTTGGCAACGAAATGAAGAACATATGGGCAAAAGACTGCTGT	514
QY	415	CAGATAGAAGTCACGCTCAGAAATTTTCTCAAGGTAGAGAGAGGCTGAAGCTAAA	474
DB	515	CAGATCAGAAGTCATGCACAAAGTCTTACAAAGTTGGAGAAAGGCTGAAGTTAAA	574
QY	475	GGTGTAGCTATGGTCAAGCGCTAGACATAGCTATTCCTCTCCACCGCCTAAGCGTAAA	534
DB	575	GGCATCCCTGTTGCCAAGCTTTGGACATAGAAATTCGGCTCCTCGCTTAAACGAAA	634
QY	535	CCAAACAAATCCTTATCCTCGAAAGACGGGAAGTGGAAACGATTCCTTATGTCAAAAACGGGT	594
DB	635	CCCAATACTCTTATCTCTGAAAACCTGGGAACAAACCGGTACATCTTCTCTCAAGTATCA	694
QY	595	GTGAATGATGAAAAGATCCCTTGGATCAGAAAAGTGTGGATCTGAGATGGCCAAT	654
DB	695	TCAGCAAAAGATGCAAACTTGTTTTCATCGGCTCTTCTTTCACAGTTGAATCAGGGGTC	754
QY	655	GAAGATCGCAACAAATCAAGCCTGAAAGAAAAC-----TCTGCAGGAA	699
DB	755	TTGGAATTGAAAATAATGCCGTTCTCTGAGAAAACATCACTGGAAAAGAAATCAGAT	814
QY	700	GACAACTGTTTCAAGATGTTTCTACATCATGATCTCTCTCTGCTGATCCTCCATGAATAAA	759
DB	815	GAGAAATGCTCGGGGTGTTTCTACTGTGAAACAAAGTATCCCTTACCAACGAAACAGGTAAGT	874
QY	760	AGTTGTATAGACATCAACGCAAGCACATTT-----CCGCGAGTTCTTGCTTC	809
DB	875	GGCGACATTGAAACAAAGTAAGACCTCACTGTGGACAAACCGGTTCAAGATGTTCCCAAG	934
QY	810	ACGGGAAGGGAGGTGAGAAATTAACAGGTA-----AGAAAAGGAGTCAAAAC	855
DB	935	AAGAACAAAGACAAAGATGTAACGATGTTACTGTGACAGCATGCAAACTACCCT	994

QY	856	TCAGATTGTAATGCAAAATCTCTGGAACCGTAATGAGCAAGACCTCAGACTTATCC-	914
DB	995	TGGCATTTCCACGCGAGATATTGTGAACGGGAATAGCAAAATGCCCTCAAAATCATCC	1054
QY	915	-----G	915
DB	1055	TCAGGTATGTTATCTCAAGACTTCATGTTTTCATCTATGAGAGAAGAAACTCACGGCAC	1114
QY	916	ATGCATATCCTGTGCTAGTGCCATTCGGGAGCTCAATAACAAGTTCTCTATCACATCCT	975
DB	1115	GCAATCTTCAAGCTACACAGCATCTGTACTACTACAGCTTCTCATCAAGCGTTTCCA	1174
QY	976	CCTTCAGAGCCAGATAGTATCCCCACACAGTTTCAGGAGATTATCAGTCGTTTCTTAAT	1035
DB	1175	GCTTGTCATTTCACAGGATGATTACCGTTTCGTTTCCAGATATCATCTACTTCTCCAAT	1234
QY	1036	CATATAATGTCACACCTTTTACAAACACCGCTCTTTATATCTGCGGCAACTTTCGCTCA	1095
DB	1235	CTTATTATGTCACCTCTCTACAGAACTCTGCAGCTCATGCTGCAGCTACATTCGCTGCT	1294
QY	1096	TCATTTTGGCTCCCG-----ATTCTAGTGGTGGCTCACCTGTTCCAGGGAAAC	1143
DB	1295	TCGCTCTGGCTTATGCGAGTGTGGGAATTCGTGATTCATCAACCCCAATGAGCTCT	1354
QY	1144	TCACCTCCGAATCTGGCTGCCATGGCCGCGACCACTGTTGCAGCTGTCTAGTCTTGGTGG	1203
DB	1355	TCTCTCCAAAGTATACTGCCATTCGCCGTGTCTACAGTAGCTGTGCAACTGCTTGGTGG	1414
QY	1204	GCTGCCAATGGATTATACCTTTATGTGCTCTCTTATGTCAGGTGTTTCACTAGTCTAT	1263
DB	1415	GCTTCTCATGGACTTCTTCTGTATGCGCTCCAGCTCCAATAACATGTGTTCCATTTCTCA	1474
QY	1264	CCTCCATCTACTTTTGGACCATCATGTGATCTAGAGTACACAAAGCAAGCATTTTACAA	1323
DB	1475	ACTGTTGCAAGTCCCACTCCAGCAATGACTGAAATGGATACCGTTGAAAATACTCAACCG	1534
QY	1324	CATGGTTCTCTGCGAGCGCAGAGCAAGAACACTCCGAGGCATCAAAAGGCTCGATCTTCA	1383
DB	1535	TTTGAGAAAACAAACACAGCTCTGCAAGATCAAACTTGGCTTCAAAATCTCCAGCTTCA	1594
QY	1384	CTGACTCAGAGGATTTGAAAT-----AAGAGTAAACAGTTTGTCTAGTAGCAG	1434
DB	1595	TCATCTGATGATTGATGAGACTGGAGTAAACCAAGCTAAATGCGACTCAAAAACCAAT	1654
QY	1435	CCTTCTGCAACCTGAGAGTGATGCAA-----AGGTTTCAGATGGAGCAGGA	1482
DB	1655	GATGATAAAATGAGGAGTTGTTTACTGCGCTGTGATGACTCAAACTGCGCCAG	1714
QY	1483	GACAGAAAACAAAGTTGACCGTCTCTGTTGGCTCAAAACACTCCGTCGAGTAGTGATGAT	1542
DB	1715	AAGAAAATCTTGTGGACCGCTCATCGTGTGGCTCAAAATACACCTTCAGGAGTGACGCA	1774
QY	1543	GTTGAGCGGATGATCAGAAAAGCAAGAGATGGCACCAATGTGTAGGTGAAAGAAACG	1602
DB	1775	GAAACTGATGATTTAGATAAAATGGAGAAAGATAAGAGGATGTGAAGGAGACAGATGAG	1834
QY	1603	AAT-----GAAGACACTATAAACCTCAAACTTCAGAGTCCATGCAACCGCGCAGTAGA	1656
DB	1835	AATCAGCGATGTTTATTTAGTTAAATTAACCGTAAGATTAAATGAGAGACAAACACAGC	1894
QY	1657	ATCAGCTCCAATATAACCGATCCATGGAAGTCTGTGTCTGACAGGGGTGCAATTTGCCTTC	1716
DB	1895	AACAACTTCACTACTGATTCGTGGAAGAAAGTCTCCGAGAGGGTCTGTATAGCGTTT	1954
QY	1717	CAAGCTCTCTTCTCAGAGAGGATTGTCGCCCAAGTTTATATATCAGAGAGAAACACAGA	1776
DB	1955	CAGGCTCTCTTTCAGAGAGAAAGATTGCTTCAAAAGCTTTTTCGCTCTCCTCAAGTGGCAG	2014
QY	1777	GAGGAGAGAACAAACAAACAGACAAAGATATCCATGTCACCTTGTATCTTAATCTTCA	1836
DB	2015	AATGTAATGAGAAAACAAAGTGACAGT-----CAATGCCATTGGCTCTTAATTTCAA	2068
QY	1837	GCTCAGTTAACACCGAGTTGATGATCAAGAGGAGAGAGAAACACAGGATTTCTTGGAAATC	1896





KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
OS Arabidopsis thaliana.

PN EP1033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128214.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131149.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141267.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2003, 01:20:08 ; Search time 4438.31 Seconds  
(without alignments)  
12343.052 Million cell updates/sec

Title: US-10-084-553-3

Perfect score: 2254

Sequence: 1 tgaatttccttcatttcgtt.....tttcccaaaaaaaaaaaaaa 2254

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	721.8	32.0	740	28	AQ958316 LERAW81TF
2	485.4	21.5	487	9	AI992931
C 3	425	18.9	426	9	AV804753
C 4	422.2	18.7	439	9	AV792310

5	395.8	17.6	515	9	AA394931	AA394931	26812	Lam
C 6	393	17.4	393	9	AV814102	AV814102	AV814102	
C 7	384	17.0	396	9	AV801368	AV801368	AV801368	
8	378.4	16.8	628	9	AV823536	AV823536	AV823536	
9	369.2	16.4	480	14	T41938	T41938	5201	Lambda
10	366.4	16.3	393	28	AQ958317	AQ958317	LERAW81TF	
C 11	360	16.0	360	9	AV784525	AV784525	AV784525	
C 12	356.2	15.8	423	9	AV807633	AV807633	AV807633	
C 13	315.6	14.0	403	9	AV799191	AV799191	AV799191	
14	292.4	13.0	334	9	AV831875	AV831875	AV831875	
15	292.4	13.0	580	9	AV827752	AV827752	AV827752	
C 16	240.8	10.7	396	9	AV818056	AV818056	AV818056	
17	215.6	9.6	805	14	CB293827	CB293827	UCRCS01	
18	206.8	9.2	771	13	BU868664	BU868664	M118F07_P	
C 19	203	9.0	511	10	BE331563	BE331563	sp15D08.Y	
C 20	198.8	8.8	560	10	BG156656	BG156656	8ab31d11.	
C 21	195.6	8.7	202	9	AV800550	AV800550	AV800550	
22	195.6	8.7	586	10	AW979367	AW979367	EST310415	
23	182.6	8.1	604	10	BG522415	BG522415	21-23 Ste	
24	180.4	8.0	1466	11	AV103618	AV103618	Zea mays	
25	180.4	8.0	562	12	BQ048428	BQ048428	1091008A0	
26	179.2	8.0	752	14	CB618876	CB618876	OSIIEa03B	
27	179.2	8.0	767	14	CB618834	CB618834	OSIIEa03A	
28	178.2	7.9	551	12	BM322287	BM322287	PtCl 2_F0	
29	177.6	7.9	656	14	CB634266	CB634266	OSIIEb13K	
30	177.6	7.9	786	14	CB671268	CB671268	OSJNEe04M	
31	177.6	7.9	796	14	CB648689	CB648689	OSJNEb12D	
32	177.6	7.9	797	14	CB629822	CB629822	OSIIEb06F	
33	177.6	7.9	800	14	CB673163	CB673163	OSJNEe07J	
34	177.6	7.9	805	14	CB633521	CB633521	OSIIEb12G	
35	177.6	7.9	811	14	CB627103	CB627103	OSIIEb01N	
36	176.8	7.9	865	14	CB635701	CB635701	OSIIEb16F	
37	176.8	7.8	595	10	BG524104	BG524104	38-82 Ste	
38	175.6	7.8	587	14	CD056977	CD056977	Hol3M22S	
39	175	7.8	808	14	CB630014	CB630014	OSIIEb06L	
40	169.4	7.5	604	10	BG524065	BG524065	38-47 Ste	
41	164.8	7.3	640	10	BG524808	BG524808	7-63 Stev	
42	164	7.3	516	9	AI900554	AI900554	sc07g01.Y	
43	159.6	7.1	672	14	CB645807	CB645807	OSJNEb07M	
C 44	156	6.9	156	9	AV562760	AV562760	AV562760	
45	155.2	6.9	612	10	BG525538	BG525538	50-93 Ste	

## ALIGNMENTS

RESULT 1  
AQ958316/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AQ958316 740 bp DNA linear GSS 28-JAN-2000  
LERAW81TF LERA Arabidopsis thaliana genomic clone LERAW81, genomic  
survey sequence.  
AQ958316 GI:6786017  
AQ958316.1  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 740)  
Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T.,  
Felblyum, T., Liang, F., Creasy, T. and Fraser, C.M.  
Genomic survey sequencing of Landsberg erecta ecotype of  
Arabidopsis thaliana and identification of sequence-based  
polymorphisms  
Unpublished  
Contact: Xiaoying Lin  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: at@tigr.org  
For additional information, see <http://www.tigr.org/tdb/at/at.html>

```
Seq primer: TF
Class: shotgun.
Location/Qualifiers
1..740
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Landsberg erecta"
/db_xref="taxon:3702"
/clone_lib="LERAW81"
/notes="Organ: Leaf; Vector: pHOS1; Total genomic DNA was
sheared to 0.9-1 Kbp before ligation."
BASE COUNT 184 a 176 c 178 g 202 t
ORIGIN
Query Match 32.0%; Score 721.8; DB 28; Length 740;
Best Local Similarity 99.0%; Pred. No. 1.1e-113;
Matches 726; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 978 TTGAGAGCCAGATAGTATCCACACAGTTGCAGGAGATTATCAGTCGTTTCTTAATCA 1037
|
|
|
Db 740 TTGAGAGCCAGATAGTATCCACACAGTTGCAGGAGATTATCAGTCGTTTCTTAATCA 681
|
|
|
QY 1038 TATAATGCTCAACCTTTTACAAACACCGCTCTTTATCTGCGCACTTTCGCTCATC 1097
|
|
|
Db 680 TATAATGCTCAACCTTTTACAAACACCGCTCTTTATCTGCGCACTTTCGCTCATC 621
|
|
|
QY 1098 ATTTGGCCTCCCGATTTAGTGTGCTCACCTGTTCCAGGGAATCTCACCTCCGAATCT 1157
|
|
|
Db 620 ATTTGGCCTCCCGATTTAGTGTGCTCACCTGTTCCAGGGAATCTCACCTCCGAATCT 561
|
|
|
QY 1158 GGTGCGCATGGCGCGAGCCACTGTTGCAGCTGCTAGTGTGTTGGTGGCTGCCAATGATT 1217
|
|
|
Db 560 GGTGCGCATGGCGCGAGCCACTGTTGCAGCTGCTAGTGTGTTGGTGGCTGCCAATGATT 501
|
|
|
QY 1218 ATTACCTTTATGTGCTCTCTTGTAGTTCAGTGGTTTTCAGTAGTCATCTCCATCTACTTT 1277
|
|
|
Db 500 GTTACCTTTATGTGCTCTCTTGTAGTTCAGTGGTTTTCAGTAGTCATCTCCATCTACTTT 441
|
|
|
QY 1278 TGGACCATCATGTGATGTAGAGTACACAAAGCAAGCAGCTTTTACACATGGTTCTGTGCA 1337
|
|
|
Db 440 TGGACCATCATGTGATGTAGAGTACACAAAGCAAGCAGCTTTTACACATGGTTCTGTGCA 381
|
|
|
QY 1338 GAGCCGAGACAGAACACTCCGAGGCATCAAAAGGCTCGATCTTCACTGGACTCAGAGGA 1397
|
|
|
Db 380 GAGCCGAGACAGAACACTCCGAGGCATCAAAAGGCTCGATCTTCACTGGACTCAGAGGA 321
|
|
|
QY 1398 TGTGAAAAATAAGAGTAAACCCAGTTTGTATGAGCAGCCTTCTGCAACACTTCAGAGTGA 1457
|
|
|
Db 320 TGTGAAAAATAAGAGTAAACCCAGTTTGTATGAGCAGCCTTCTGCAACACTTCAGAGTGA 261
|
|
|
QY 1458 TGCAAAAGGTTTCAGATGGAGCAGAGACAGAAACCAAGTTGACCGTCTCTGTTGGCTC 1517
|
|
|
Db 260 TGCAAAAGGTTTCAGATGGAGCAGAGACAGAAACCAAGTTGACCGTCTCTGTTGGCTC 201
|
|
|
QY 1518 AATCACTCCGTCCAGTAGTATGATCTTGGAGCGGATGATCATGAAGGCAAGGATGG 1577
|
|
|
Db 200 AATCACTCCGTCCAGTAGTATGATCTTGGAGCGGATGATCATGAAGGCAAGGATGG 141
|
|
|
QY 1578 CACCAATGGTGGAGTGAAGAAACGAATGAAGACACTAATAAACCCTCAAACTTCAGAGTC 1637
|
|
|
Db 140 CTCCAATGGTGGAGTGAAGAAACGAATGAAGACACTAATAAACCCTCAAACTTCAGAGTC 81
|
|
|
QY 1638 CAATGACGCGCAGTAGAATCAGCTCCCAATATAACCGATCCCATGGAAGTCTGTGCTGA 1697
|
|
|
Db 80 CAATGACGCGCAGTAGAATCAGCTCCCAATATAACCGATCCCATGGAAGTCTGTGCTGA 21
|
|
|
QY 1698 CGAGGGTCGAATT 1710
|
|
|
Db 20 CGAGGGTACTTACT 8
|
|
|
RESULT 2
```

```
AI992931
LOCUS
DEFINITION
701495069 A. thaliana, Ohio State clone set Arabidopsis thaliana
cdna clone 701495069, mRNA sequence.
AI992931
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 487)
Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R.,
Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Carpio, I.,
Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrigha, A., Murry, L.,
Turner, C., Krikorian, S., Elder, L. and Hanson, D.
Arabidopsis thaliana Gene Expression MicroArray
Unpublished
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.
Location/Qualifiers
1..487
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="701495069"
/clone_lib="A. thaliana, Ohio State clone set"
/notes="cDNA library was made from selected clones from the
Arabidopsis thaliana Ohio State clone set."
BASE COUNT 160 a 105 c 115 g 107 t
ORIGIN
Query Match 21.5%; Score 485.4; DB 9; Length 487;
Best Local Similarity 99.8%; Pred. No. 3.7e-73;
Matches 486; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 466 GAAGCTAAAGGTGTAGCTATGGGTCAAGCGCTAGACATAGCTATTCCTCTCCACGCGCT 525
|
|
|
Db 1 GAAGCTAAAGGTGTAGCTATGGGTCAAGCGCTAGACATAGCTATTCCTCTCCACGCGCT 60
|
|
|
QY 526 AAGCGTAAACCAACAACTCTTATCTCGAAAGACGGGAAGTGGACGATCCTTATGTCA 585
|
|
|
Db 61 AAGCGTAAACCAACAACTCTTATCTCGAAAGACGGGAAGTGGACGATCCTTATGTCA 120
|
|
|
QY 586 AAAACCGGTGTGAATCATGGAAGAGTCCCTTGGATCAGAAAAGTGTCCGATCCTGAG 645
|
|
|
Db 121 AAAACCGGTGTGAATCATGGAAGAGTCCCTTGGATCAGAAAAGTGTCCGATCCTGAG 180
|
|
|
QY 646 ATGGCCAAATGAAGATCGAACAACTCAAGCCCTGAAGAGAAAACCTCTGCAAGGAGACAAC 705
|
|
|
Db 181 ATGGCCAAATGAAGATCGAACAACTCAAGCCCTGAAGAGAAAACCTCTGCAAGGAGACAAC 240
|
|
|
QY 706 TGTTCAGATGTTTCTACTCATCAGTATCTCTGCTGCATCCTCCATGAATAAAGTTGT 765
|
|
|
Db 241 TGTTCAGATGTTTCTACTCATCAGTATCTCTGCTGCATCCTCCATGAATAAAGTTGT 300
|
|
|
QY 766 ATAGAGACATCAAAACCAAGCACTTTCCGAGGTTCTTGGCTTTCACGGGAAGGGAAGT 825
|
|
|
Db 301 ATAGAGACATCAAAACCAAGCACTTTCCGAGGTTCTTGGCTTTCACGGGAAGGGAAGT 360
|
|
|
QY 826 CAGAATAACAGGGGTAGAAAAGGAGTCAAACTCAGATTTGAATGAAAATCTCTGGAAAAC 885
|
|
|
Db 361 CAGAATAACAGGGGTAGAAAAGGAGTCAAACTCAGATTTGAATGAAAATCTCTGGAAAAC 420
|
|
|
QY 886 GGTAATGAGCAAGGAGCTCAGACTTTATCCGATGCATATCCCTGTGTGTAGTGCCATTGGG 945
```

```

|||||
421 GGTAATGACAAAGGACCTCAGACTTATCCGATGCATATCCCTGTGTAGTGCAATGGGG 480
|||||
QY 946 AGCTCAA 952
|||||
Db 481 GGCTCAA 487
|||||

RESULT 3
AV804753/c
LOCUS AV804753 RAFL9 Arabidopsis thaliana cdna clone RAFL09-40-G07 3', EST 29-MAR-2002
DEFINITION mRNA sequence.
ACCESSION AV804753
VERSION AV804753.1 GI:19838738
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 426)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cdna (2002b)
Unpublished
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cdna library was constructed essentially
as reported previously (Seki et al., 1998). cdna cleaved with BamHI
and XhoI was ligated to modified lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
Location/Qualifiers
1..426
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-40-G07"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
136 a 75 c 82 g 133 t

BASE COUNT
ORIGIN

Query Match 18.9%; Score 425; DB 9; Length 426;
Best Local Similarity 100.0%; Pred. No. 8.4e-63;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1804 AGATATCCAATGGCACTTGATCTTAACCTTCACAGCTCAGTTAAACACCGATTCATGATCAA 1863
|||||
Db 426 AGATATCCAATGGCACTTGATCTTAACCTTCACAGCTCAGTTAAACACCGATTCATGATCAA 367
|||||

QY 1864 GAGGAGAAGAGAAACACAGGATTTCTTGGAAATCGGATTAGATGCTTCAAGCTATGAGT 1923
|||||
Db 366 GAGGAGAAGAGAAACACAGGATTTCTTGGAAATCGGATTAGATGCTTCAAGCTATGAGT 307
|||||

QY 1924 AGAGGAAGACAGGTTTTTAAACCATACAAAGATGTTTCATGGAAGCCAAAGAAAGTAGA 1983
|||||
Db 306 AGAGGAAGACAGGTTTTTAAACCATACAAAGATGTTTCATGGAAGCCAAAGAAAGTAGA 247
|||||

```

```

QY 1984 ATCTCTCAACAACAATCCTATCATTTCTGTGGAAACAGAAAGATCCAAACGGATCGCGTTG 2043
|||||
Db 246 ATCTCTCAACAACAATCCTATCATTTCTGTGGAAACAGAAAGATCCAAACGGATCGCGTTG 187
|||||

QY 2044 GAAACTCAAGCTTCCACATGAGACTCTATTTTCATCTGATCTGTGTTGTTGACTCTGTTT 2103
|||||
Db 186 GAAACTCAAGCTTCCACATGAGACTCTATTTTCATCTGATCTGTGTTGTTGACTCTGTTT 127
|||||

QY 2104 TTAAGTTTTTCAAGACCACTGCTACATTTCTTTTCTTTTGGGCGCTTTCTATTGTTTC 2163
|||||
Db 126 TTAAGTTTTTCAAGACCACTGCTACATTTCTTTTCTTTTGGGCGCTTTCTATTGTTTC 67
|||||

QY 2164 CTTGTCCATAGTCTTCTCTGAACATTTGACTCTGTATTATTCAACAAATCATAACTGTT 2223
|||||
Db 66 CTTGTCCATAGTCTTCTCTGAACATTTGACTCTGTATTATTCAACAAATCATAACTGTT 7
|||||

QY 2224 TAATC 2228
Db 6 TAATC 2

RESULT 4
AV792310/c
LOCUS AV792310 RAFL7 Arabidopsis thaliana cdna clone RAFL07-14-G14 3', EST 29-MAR-2002
DEFINITION mRNA sequence.
ACCESSION AV792310
VERSION AV792310.1 GI:19826293
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 439)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cdna (2002b)
Unpublished
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cdna library was constructed essentially
as reported previously (Seki et al., 1998). cdna cleaved with BamHI
and XhoI was ligated to modified lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
Location/Qualifiers
1..439
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL07-14-G14"
/dev_stage="rosette plants"
/lab_host="DH10B"
/clone_lib="RAFL7"
/note="Site 1: BamHI; Site 2: SalI; subjected to
cold-treated (1, 2, 5, 10, 24 hr)"
144 a 74 c 86 g 135 t

BASE COUNT
ORIGIN

Query Match 18.7%; Score 422.2; DB 9; Length 439;
Best Local Similarity 98.2%; Pred. No. 2.5e-62;
Matches 427; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

/note=Vector: lambda Zip-Lox; Site\_1: Sal; Site\_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light; 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA. "

Query Match	17.6%	Score 395.8	DB 9	Length 515
Best Local Similarity	91.9%	Pred. No. 8e-58		
Matches 431	Conservative 0	Mismatches 36	Indels 2	Gaps 2
Qy 1770	ACACAGAGAGGAGAACCAACAAACAAAGAACAAAGATATCCAAATGGCACTTGCATCTTAA	1829		
Db				
Qy 49	ACCCCGGGGGAGGTCCACCNACNCGACCCAGNNTTCCCAAGGCGNCTTG-TCTTAA	107		
Db				
Qy 1830	CTTTCACAGCTCAGTTAAACACCCAGTTCGATGATCAAGAGGAGAGACGAGAAACACAGGATTTCT	1889		
Db				
Qy 108	CTTCNCGGCCAGTTAACCCCGTGTGTGACCCAGGGGNGAGAGANNWCAGA-TTCT	166		
Db				
Qy 1890	TGGAATCGGATTAGATGCTTTCAAAGCTTAATGAGTAGAGGAAGAACAGGTTTTTAAACCATA	1949		
Db				
Qy 167	TGGAANCGGATTAGATGCTNCAAGCTTAATGAGTAGAGGAAGAACAGGTTTTTAAACCATA	226		
Db				
Qy 1950	CAAAAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCCTCAACAACAATCCTATCATTC	2009		
Db				
Qy 227	CAAAAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCCTCAACAACAATCCTATCATTC	286		
Db				
Qy 2010	TGTGGACACAGAAAGATCCCAACCGGATGCGGTTGGAAACTCAAGCTTCCACATGAGACTC	2069		
Db				
Qy 287	TGTGGACACAGAAAGATCCCAACCGGATGCGGTTGGAAACTCAAGCTTCCACATGAGACTC	346		
Db				
Qy 2070	TATTTTCATCTGATCTGTTGTTGTACTCTGTTTTTAAAGTTTTTCAAGACCACATGCTACAT	2129		
Db				
Qy 347	TATTTTCATCTGATCTGTTGTTGTACTCTGTTTTTAAAGTTTTTCAAGACCACATGCTACAT	406		
Db				
Qy 2130	TTTCTTTTTCTTTTGAGGCCCTTTGTATTTGTTTCTTGTGTCATAGTCTTCTCTGTAACATT	2189		
Db				
Qy 407	TTTCTTTTTCTTTTGAGGCCCTTTGTATTTGTTTCTTGTGTCATAGTCTTCTCTGTAACATT	466		
Db				
Qy 2190	TGACTCTGTATTATTCAACAAATCATAAACTGTTTAACTCTTTTTTTTC	2238		
Db				
Qy 467	TGACTCTGTATTATTCAACAAATCATAAACTGTTTAACTCTTTTTTTTC	515		
Db				

RESULT 6				
AV814102/c				
LOCUS	AV814102	393 bp	mRNA	linear EST 01-APR-2002
DEFINITION	AV814102	RAFL9	Arabidopsis thaliana cDNA clone	RAFL09-81-L16 3', mRNA sequence.

ACCESSION	AV814102	
VERSION	AV814102.1	GI:19855881
KEYWORDS	EST.	
SOURCE	Arabidopsis thaliana	(thale cress)
ORGANISM	Arabidopsis thaliana	
	Eukaryota; Viridiplantae;	Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta;	eudicotyledons; core eudicots; rosids
	; eurosid II; Brassicales;	Brassicaceae; Arabidopsais.

REFERENCE  
1 (bases 1 to 393)  
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
Large scale analysis of Arabidopsis full-length cDNA (2002b)  
Unpublished  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
TITLE  
JOURNAL  
COMMENT

RESULT 5  
AA394931  
LOCUS  
DEFINITION  
26812 Lambda-PRL2 Arabidopsis thaliana cDNA clone 109K9XP 3', mRNA  
515 bp linear EST 30-OCT-1997  
AA394931  
SEQUENCE  
26812 Lambda-PRL2 Arabidopsis thaliana cDNA clone 109K9XP 3', mRNA  
515 bp linear EST 30-OCT-1997  
AA394931  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids  
; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 515)  
NEWMAN,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh  
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel  
,E. and Somerville,C.  
Genes Galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
Plant Physiol. 106, 1241-1255 (1994)  
95148729  
7846151  
Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.

```

Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ibm.cl.msu.edu
Seq primer: M13 -21 dye primer.
Location/Qualifiers
1. 515
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="var columbia"
/db xref="taxon:3702"
/clone="109K9p"
/clone_lib="Lambda-pb12"
FEATURES
source

```



DEFINITION AV823536 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-20-N17 5', mRNA sequence.

ACCESSION AV823536

VERSION AV823536.1 GI:19865596

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 628)

AUTHORS Seki M., Narusaka M., Ishida J., Kamiya A., Satou M., Nakajima M., Oono Y., Sakurai T., Garrinci P., Kawai J., Itoh M., Ishii Y., Arakawa T., Shibata K., Shingawa A., Muramatsu M., Hayashizaki Y., and Shinozaki K.

TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)

JOURNAL Unpublished

COMMENT Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: mseki@rtc.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a Sati/XhoI insert. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

FEATURES	source	Location/Qualifiers
1. .628		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/db_xref="taxon:3702"
		/clone="RAFL05-20-N17"
		/dev_stage="rosette plants"
		/lab_host="SOLR"
		/clone_idb="RAFL5"
		/note="Site 1: SstI; Site 2: XhoI"
194 a	94 c	dehydration-treated(1,2,5,10,24
BASE COUNT	194 a	127 g 210 t

Query Match	16.8%;	Score	378.4;	DB	9;	Length	628;
Best Local Similarity	99.5%;	Pred. No.	7.2e-55;				
Matches	390;	Conservative	0;	Mismatches	1;	Indels	1;
QY	65	ACTCTCTCTCTCTCTCTCTCTCGATTCTTACTGTTTCTTATCCCAACGAAATCTCGAAT	124				
Db	2	ACTCTCTCTCTCTCTCTCTCTCGATTCTTACTGTTTCTTATCCCAACGAAATCTCGAAT	61				
QY	125	TAAAAATGGAATCTTTTATTCGAATCAAG-CTGATTTTGGTTTCTTTCATTTGAATCATCTCT	183				
Db	62	TAAAAATGGAATCTTTTATTCGAATCAAGCCTGATTTTGGTTTCTTTCATTTGAATCATCTCT	121				
QY	184	CTAAAGTGGAAATTTTGTAAAGAGAAGATCTCGAAGTTGTGTAGAGGAGCTTTAGTGATGGAG	243				
Db	122	CTAAAGTGGAAATTTTGTAAAGAGAAGATCTCGAAGTTGTGTAGAGGAGCTTTAGTGATGGAG	181				
QY	244	ACAAATTCGTCTGGAGAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATAACAAAG	303				
Db	182	ACAAATTCGTCTGGAGAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATAACAAAG	241				
QY	304	CAACCTGGAAGGTGGACTCTGAGAAGACATAATAGATTTCATTGAAGCTTTGAGGCTTTAT	363				
Db	242	CAACCTGGAAGGTGGACTCTGAGAAGACATAATAGATTTCATTGAAGCTTTGAGGCTTTAT	301				
QY	364	GGTAGAGCATGGCAGAAGATTGAAGAACATGTAGCAACAAAACACTGCTCTCCAGATAAGA	423				
Db	302	GGTAGAGCATGGCAGAAGATTGAAGAACATGTAGCAACAAAACACTGCTCTCCAGATAAGA	361				
QY	424	AGTCACGCTCAGAAATTTTCTCCAAAGGTAGA	455				

DB	362	AGTCACGCTCAGAAATTTTCTCCAAGGTAAA	393
RESULT 9			
T41938			
LOCUS	480 bp	mrna	linear
DEFINITION	5201 Lambda-PRL2 Arabidopsis thaliana	CDNA clone	109K917, mRNA
ACCESSION	T41938		
VERSION	T41938.1	GI:634526	
KEYWORDS	EST.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids I; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
AUTHORS	1 (bases 1 to 480)		
	Newman,T., deBruijn,J., Green,P., Keegstra,K., Kende,H., McIntosh L., Ohlridge,J., Raikheil,N., Somerville,S., Thomashow,M., Retzel E. and Somerville,C.		
TITLE	Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones		
JOURNAL	Plant Physiol.	106,	1241-1255 (1994)
MEDLINE	95148729		
PUBMED	7846151		
COMMENT	Contact: Thomas Newman		
	MSU-DOE Plant Research Laboratory		
	Michigan State University		
	MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.		

FEATURES  
SOURCE

```

seq primer: 171
Location/Qualifiers
1..480
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="var.columbia"
/db_xref="taxon:3702"
/clone="109K97"
/clone.lib="Lambda-PRL2"
/note="vector: lambda Zip-Lox; S
Lambda PRL2 is a cDNA library de
quantities of 4 pools of mRNA.
day germinated etiolated seedlin
grown roots; 3) staged plants ha
cycle, half on 16 hr light, 8 ho
same plants as 3 but aerial tiss
siliques. The vector is BRL's l
inserts were directionally clone
oligo dt primed cDNA."

```

BASE COUNTRY OF ORIGIN

	Query Match	16.4%	Score 369.2	DB 14	Length 480
	Best Local Similarity	94.0%	Pred. No. 2.9e-53		
	Matches 451	Conservative 0	Mismatches 22	Indels 7	Gaps 7
QY	466	GAAGCTAAAGGTGTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCT	525		
Db	1	GNAGCTTAAAGGTGTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCT	60		
QY	526	AAGCGTAAACCAACAAATCCTTATCCTCGAAGACGGGAGTGGAAACGATCCTTATGTCA	585		
Db	61	AAGCGTAAACCAACAAATCCTTATCCTCGAAGACGGGAGTGGAAACGATCCTTATGTCA	120		
QY	586	AAAAACGGGTGTGAATGTGGAAAAAGATGCCCTTGGATCAGAAAAAGTGTGCATCCTGAG	645		
Db	121	AAAAACGGGTGTGAATGTGGAAAAAGATGCCCTTGGATCAGAAAAAGTGTGCATCCTGAG	180		
QY	646	ATGGCCAAATGAAGATCGACAACAATCAAGAGCCTGAAGAGAAAAATCTTCGAGGAAGACAAC	705		

Db 181 ATGGCCATGAAGATCGACAACAATCAAGACCTGAAGAGAAACACTTCGAGGAAGACAAC 240  
 QY 706 TGCTCAGATTGTTTCACTCATCATGATATCTCTCTGCTGCATCTCTCCATGAATAAAGATTGCT 765  
 Db 241 TTTTCAGATTGTTTCACTCATCATGATATCTCTCTGCTGCATCTCTCCATGAATAAAGATTGCT 300  
 QY 766 ATAGAGACATCAAA-CGCAAGACATTTCCGCGAG-TTCTTGCTCTTCCGCGAA-GAGGGA 822  
 Db 301 ATAGNGCATCAAAACCGCAAGACATTTCCGCGAGTTCTTGNCTTCAACGGAAGGAGGA 360  
 QY 823 ACTCAGATAACAGGTAAG-AAAGGAGTCAACATCAGATTTCAATGC-AAATCTCTCG 880  
 Db 361 ACTCAGNATAACAGGTAAGAAAGAGTCAACATCAGGTTTGAATTCAAAATCTCTG 420  
 QY 881 AAAACGGTAA-TCAGCAAGGACCTCAGACTTATCC-GATGCATATCCCTGTGCTAGTGCC 938  
 Db 421 AAAACGGTAATTNGCAAGNCTCAGACTTTTNCGGTTGAATATCCCTGTGCGAGTGCC 480

RESULT 10  
 AQ958317  
 LOCUS LERAW81TR LERA Arabidopsis thaliana genomic clone LERAW81, genomic survey sequence.  
 DEFINITION  
 ACCESSION AQ958317.1 GI:6786018  
 VERSION  
 KEYWORDS GSS.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 393)  
 AUTHORS Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T., Fiedllyum, T., Liang, F., Creasy, T. and Fraser, C.M.  
 TITLE Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms  
 JOURNAL Unpublished  
 COMMENT Contact: Xiaoying Lin  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: at@tigr.org  
 For additional information, see <http://www.tigr.org/tdb/at/at.html>  
 Seq primer: TR  
 Class: shotgun.

FEATURES  
 source  
 1..393  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Landsberg erecta"  
 /db\_xref="taxon:3702"  
 /clone="LERAW81"  
 /clone\_lib="LERA"  
 /note="Organ: Leaf; Vector: PHOS1; Total genomic DNA was sheared to 0.9-1 kbp before ligation."  
 BASE COUNT 90 a 108 c 84 g 111 t  
 ORIGIN  
 Query Match 16.3%; Score 366.4; DB 28; Length 393;  
 Best Local Similarity 98.4%; Pred. No. 9.3e-53;  
 Matches 370; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 857 CAGATTGTAATCAAAATCTCTGGAACGTAATGAGCAAGACCTCAGACTTATCCGA 916  
 Db 13 CAGATTGTAATCAAAATCTCTGGAACGTAATGAGCAAGACCTCAGACTTATCCGA 72  
 QY 917 TGCATATCCCTGTGCTAGTGCCATTTGGGAGCTTCAATAACAAGTTCTCTATCACATCCTC 976  
 Db 73 TGCATATCCCTGTGCTAGTGCCATTTGGGAGCTTCAATAACAAGTTCTCTATCACATCCTC 132

AV784525 360 bp mRNA linear EST 28-MAR-2002  
 AV784525 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-20-N17 3', mRNA sequence.  
 ACCESSION AV784525.1 GI:19803315  
 VERSION  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 360)  
 AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
 TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 JOURNAL Unpublished  
 COMMENT Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: msekita@tc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pluescript vector as a SstI/XhoI insert. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.  
 FEATURES  
 Location/Qualifiers  
 1..360  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL05-20-N17"  
 /dev\_stage="rosette plants"  
 /lab\_host="SOLR"  
 /clone\_lib="RAFL5"  
 /note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated(1,2,5,10,24 hr)"  
 BASE COUNT 128 a 59 c 69 g 104 t  
 ORIGIN  
 Query Match 16.0%; Score 360; DB 9; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-51;  
 Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1880 CAGGATTTCCTGGAATCGGATTAGATGCTTCAAGCTAATAGTAGGAGGAACAGGTT 1939



```

Db      360 CAGGATTCTTGGAAATCGGATTAGATGCTTTCAAAGCTTAATGAGTAGAGGAAGAACAGGTT 301
QY      1940 TTAACACCATACAAAAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCCTCAACAAACATC 1999
Db      300 TTAACACCATACAAAAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCCTCAACAAACATC 241
QY      2000 CTATCATTCATGTGGAAACAGAAAGATCCCAACCGGATCGGTTGGAAACTCAAGCTTCCA 2059
Db      240 CTATCATTCATGTGGAAACAGAAAGATCCCAACCGGATCGGTTGGAAACTCAAGCTTCCA 181
QY      2060 CATGAGACTCTATTTTCATCTGATCTGTTGTTGTGACTCTGTTTAAAGTTTTCAGAAC 2119
Db      180 CATGAGACTCTATTTTCATCTGATCTGTTGTTGTGACTCTGTTTAAAGTTTTCAGAAC 121
QY      2120 ACTGCTACATTTCTTTCTTTTGAGGCCCTTTGTTATTTCTTCTTCTGTCATAGTCTTC 2179
Db      120 ACTGCTACATTTCTTTCTTTTGAGGCCCTTTGTTATTTCTTCTTCTGTCATAGTCTTC 61
QY      2180 CTGTAACATTTGACTCTGTATTATTCACAAATCATAAACTGTTTAACTTTTTTTTTC 2239
Db      60 CTGTAACATTTGACTCTGTATTATTCACAAATCATAAACTGTTTAACTTTTTTTTTC 1

RESULT 12
AV807633/c
LOCUS   AV807633 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-51-F03 3',
DEFINITION
mRNA sequence.
ACCESSION
AV807633
VERSION
AV807633.1 GI:19841618
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 423)
SEKI,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluscript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.

FEATURES
            Location/Qualifiers
                1..423
                    /organism="Arabidopsis thaliana"
                    /mol_type="mRNA"
                    /db_xref="taxon:3702"
                    /clone="RAFL09-51-F03"
                    /dev_stage="plants at various developmental stages from
                    germination to mature seeds"
                    /lab_host="DH10B"
                    /clone_lib="RAFL9"
                    /notes="Site 1: BamHI; Site 2: SalI; subjected to
                    dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                    hr) treatments"
BASE COUNT      143 a      70 c      89 g      121 t
ORIGIN

```

```

Query Match      15.8%; Score 356.2; DB 9; Length 423;
Best Local Similarity 96.0%; Pred. No. 51e-51;
Matches 408; Conservative 0; Mismatches 13; Indels 4; Gaps 4;

QY      1816 GCACCTTGATCTTAACCTTCACAGCTCAGTTAAACACCACTTGATGATCAAGAGGAGAAGA 1875
Db      423 GCACCTTGATCTTAACCTTCACAGCTCAGTT-ACACCATTTGATGATCAAGAGGAGAAGA 365
QY      1876 AACACAGGATTTCTTGGAAATCGGATTAGATGCTTCAAGCTTAATGAGTAGAGGAAGAAC 1935
Db      364 AACACAGGATTTCTTGGAAATCGGATTAGATGCTTCAAGCTTAATGATTAGTA-GAAGCACA 306
QY      1936 GGTGTTTAAACCATACAAAAGATGTTCCATGGAAGCCAAAGAAAGTAGAATCCTCAACAAAC 1995
Db      305 GGTGTTTAAACCATACAAAAGATGTTCCAT-GAAGCCAAAGAAAGTAGAATCCTCAACAA 247
QY      1996 AATCCTATCATTCATGTGGAACAGAAAGATCCCAACCGGATCGGTTGGAAACTCAAGCT 2055
Db      246 CATCCTATCATTCATGTGCAACAGGAAGATCCCAACCGGATCGGTTGGAAACTCAAGCT 187
QY      2056 TCACATGAGACTCTATTTTCATCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2115
Db      186 TCACATGAGACTCTATTTTCATCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 127
QY      2116 GACCACTGCTACATTTCTTTCTTTTGAGGCCCTTTGTTATTTCTTCTTCTGTCATAGT 2175
Db      126 GACCACTCTACATTTCTTTCTTTTGAGGCCCTTTGTTATTTCTTCTTCTGTCATAGT 67
QY      2176 CTTCTGTTAACTTTGACTCTGTATTATTCACCAATCATAAAGCTGTTTAACTCTTTT 2234
Db      66 CTTCTGTTAACTTTGACTCTGTATTATTCACCAATCATAAAGCTGTTTAACTCTTTT 7
QY      2235 TTTCC 2239
Db      6 TTTCC 2

RESULT 13
AV799191/c
LOCUS   AV799191 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-19-A10 3',
DEFINITION
mRNA sequence.
ACCESSION
AV799191
VERSION
AV799191.1 GI:1983174
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 403)
SEKI,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluscript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
details.

FEATURES
            Location/Qualifiers
                1..403

```

/organism="Arabidopsis thaliana"

/mol\_type="mrna"

/db\_xref="taxon:3702"

/clone="RAFL09-19-A10"

/dev\_stage="plants at various developmental stages from

germination to mature seeds"

/lab\_host="DH10B"

/clone\_lib="RAFL9"

/note="Site 1: BamHI; Site 2: SalI; subjected to  
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
hr) treatments"

BASE COUNT 135 a 70 c 74 g 124 t

ORIGIN

Query Match 14.0%; Score 315.6; DB 9; Length 403;

Best Local Similarity 93.1%; Pred. No. 4.6e-44;

Matches 376; Conservative 0; Mismatches 19; Indels 9; Gaps 4;

QY 1843 TTAACACCAAGTTCATCAAGAGGAGAGAGAAACACAGGATTTCTGGAAATCGGATTA 1902

Db 403 TTAACACCAAGTTCATCAAGAGGAGAGAGAAACACAGGATTTCTGGAAATCGGATTA 344

QY 1903 GATGCTTCAAGCTAATGATGAGGAGAGAGAAACACAGGTTTAAACCATACAAAGATGTTCC 1962

Db 343 GATGCTTCAAGCTAATGATGAGGAGAGAGAAACACAGGTTTAAACCATACAAAGATGTTCC 284

QY 1963 ATGGAAGCCAAAGAAAGTGAATCCTCAACAACATCCTATCATTCATGTGGAACAGAAA 2022

Db 283 ATGGAAGCCAAAGAAAGTGAATCCTCAACAACATCCTATCATTCATGTGGAACAGAAA 224

QY 2023 GATCCAAACGATGCGGTGGAACTCAAGCTCCACATGAGACGTCATTTTCATCTGA 2082

Db 223 GATCCAAACGATGCGGTGG- AACTCAAGCTCCACATGAGACGTCATTTTCATCTGA 165

QY 2083 TCTGTTGTTGTTACTGTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTT 2142

Db 164 TCTGTTGTTGTTACTG- TTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTT 106

QY 2143 TGAGGCCCTT---GTAATTGTTCTGTCATAGTCTTCCTGT---AACATTTGACTC 2195

Db 105 TGAGGCCCTTGTGAATTTGTTCCCTTGCCCAATAGCTTTCCCTGTAAACATTTGACTC 46

QY 2196 TGTATATTCAACAATCAATAACTGTTTAATCTTTTTTTTTTTTCC 2239

Db 45 TGTATATTCAACAATCAATAACTGTTTAATCTTTTTTTTTTTTCC 2

RESULT 14

AV831875

LOCUS AV831875 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-96-L16 5',

DEFINITION mRNA sequence.

ACCESSION AV831875.1 GI:19873935.

VERSION AV831875

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; euroids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 334)

AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,

Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,

Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.

and Shinozaki,K.

Large scale analysis of Arabidopsis full-length cDNA (2002b)

Unpublished

CONTACT: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: mseki@rtc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Garninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified paluescript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

FEATURES

source

1..334

/organism="Arabidopsis thaliana"

/mol\_type="mrna"

/db\_xref="taxon:3702"

/clone="RAFL09-96-L16"

/dev\_stage="plants at various developmental stages from

germination to mature seeds"

/lab\_host="DH10B"

/clone\_lib="RAFL9"

/note="Site 1: BamHI; Site 2: SalI; subjected to  
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24  
hr) treatments"

BASE COUNT 113 a 47 c 78 g 93 t 3 others

ORIGIN

Query Match 13.0%; Score 292.4; DB 9; Length 334;

Best Local Similarity 99.7%; Pred. No. 4.6e-40;

Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 TTTCTTTTCATGATCATCTCTCTAAAGTGGAAATTTTGTAAAGAGAAGATCTGAAGTTGT 221

Db 1 TTTCTTTTCATGATCATCTCTCTAAAGTGGAAATTTTGTAAAGAGAAGATCTGAAGTTGT 60

QY 222 GTAGAGGAGCTTAGTCGAGACAAATTCGTCGAGAGAGATCTGGTTATTAGACTCG 281

Db 61 GTAGAGGAGCTTAGTCGAGACAAATTCGTCGAGAGAGATCTGGTTATTAGACTCG 120

QY 282 GAAGCCATATACGATAACAAAGCAACGCTGAAAGGTGGACTGAGCAAGACATATAGATT 341

Db 121 GAAGCCATATACGATAACAAAGCAACGCTGAAAGGTGGACTGAGCAAGACATATAGATT 180

QY 342 CATTTGAAGCTTTGAGGCTTTATGGTAGACATGGCAGAGATTGAAGAACATGTAGCAAC 401

Db 181 CATTTGAAGCTTTGAGGCTTTATGGTAGACATGGCAGAGATTGAAGAACATGTAGCAAC 240

QY 402 AAAAAGCTGCTCCAGATAAGAGTCACGCTCAGAAATTTTCTCCAGGTAGA 455

Db 241 AAAAAGCTGCTCCAGATAAGAGTCACGCTCAGAAATTTTCTCCAGGTAAA 294

RESULT 15

AV827752

LOCUS AV827752

DEFINITION AV827752 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-19-A10 5',

mRNA sequence.

ACCESSION AV827752

VERSION AV827752.1 GI:19869812

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; euroids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 580)

AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,

Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,

Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.

and Shinozaki,K.

Large scale analysis of Arabidopsis full-length cDNA (2002b)

Unpublished

CONTACT: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359  
Fax: 81-298-36-9060

FEATURES	Location/Qualifiers
accus.	

Query Match 13.0%; Score 292.4; DB 9; Length 580;

Search completed: Novemb  
Job time : 4444.81 secs



361 QY TATGGTAGAGCATGGCAAGAAATTGAAGAAATAGTACCAACAAAACTGCTGTCCAGATA 420  
361 Db TATGGTAGAGCATGGCAAGAAATTGAAGAAATAGTACCAACAAAACTGCTGTCCAGATA 420  
421 QY AGAAGTCACGCTCAGAAATTTTCTCCAAAGTAGAGAAAGAGGCTGAAGCTAAAGTGTA 480  
421 Db AGAAGTCACGCTCAGAAATTTTCTCCAAAGTAGAGAAAGAGGCTGAAGCTAAAGTGTA 480  
481 QY GCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCAGCGCTAAGCGCTAAACCAAC 540  
481 Db GCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCAGCGCTAAGCGCTAAACCAAC 540  
541 QY AATCCTTTATCTCGAAAGAGCGGAAGTGAACGATCCTTATGTCAAAAACGGGTGTAAT 600  
541 Db AATCCTTTATCTCGAAAGAGCGGAAGTGAACGATCCTTATGTCAAAAACGGGTGTAAT 600  
601 QY GATGGAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCGAGATGGCCAATGAAGAT 660  
601 Db GATGGAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCGAGATGGCCAATGAAGAT 660  
661 QY CGAACCAATCAAGCGCTGAAGAGAAACTCTCGAGGAAGACACTGTTCAGATTGTTTC 720  
661 Db CGAACCAATCAAGCGCTGAAGAGAAACTCTCGAGGAAGACACTGTTCAGATTGTTTC 720  
721 QY ACTCATCAGTATCTCTGCTGTCATCCTCCATGAATAAAAGTTGTATAGACATCAAAC 780  
721 Db ACTCATCAGTATCTCTGCTGTCATCCTCCATGAATAAAAGTTGTATAGACATCAAAC 780  
781 QY GCAAGCACTTTCGCGAGATTCTGCGCTTTCGCGGAGAGGGAAGTCAAGAGGTA 840  
781 Db GCAAGCACTTTCGCGAGATTCTGCGCTTTCGCGGAGAGGGAAGTCAAGAGGTA 840  
841 QY AGAAGGAGTCAAACTCAGATTGCAATGCAAAATCTCTGGAACCGTAAATGAGCAAGGA 900  
841 Db AGAAGGAGTCAAACTCAGATTGCAATGCAAAATCTCTGGAACCGTAAATGAGCAAGGA 900  
901 QY CCTCAGACTTATCGATGCAATCCCTGTGCTAGTGCCATTTGGGGAGCTCAATAACAAGT 960  
901 Db CCTCAGACTTATCGATGCAATCCCTGTGCTAGTGCCATTTGGGGAGCTCAATAACAAGT 960  
961 QY TCTCTATCATCCTCTCAGAGCCAGATAGTATCATCCACACAGTTGAGGAGATTAT 1020  
961 Db TCTCTATCATCCTCTCAGAGCCAGATAGTATCATCCACACAGTTGAGGAGATTAT 1020  
1021 QY CAGTGGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTATCTGCC 1080  
1021 Db CAGTGGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTATCTGCC 1080  
1081 QY GCAACTTTCGCTCATCATTTTGGCTCCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGG 1140  
1081 Db GCAACTTTCGCTCATCATTTTGGCTCCCGATTTCTAGTGGTGGCTCACCTGTTCCAGGG 1140  
1141 QY AACTCACCTCCGAATCTGGCTGCCATGGCCGACGCACTGTCGAGCTGCTAGTCTGG 1200  
1141 Db AACTCACCTCCGAATCTGGCTGCCATGGCCGACGCACTGTCGAGCTGCTAGTCTGG 1200  
1201 QY TGGGCTGCCAATGGATTATTAACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGT 1260  
1201 Db TGGGCTGCCAATGGATTATTAACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGT 1260  
1261 QY CATCTCCATCTACTTTTGGACCATCATGTGTAGTAGTACACAAAGCAAGCACTTTA 1320  
1261 Db CATCTCCATCTACTTTTGGACCATCATGTGTAGTAGTACACAAAGCAAGCACTTTA 1320  
1321 QY CAACATGGTTCTGTGACAGCCGAGACAGAACACTCCGAGCATCAAAAGGCTCGATCT 1380  
1321 Db CAACATGGTTCTGTGACAGCCGAGACAGAACACTCCGAGCATCAAAAGGCTCGATCT 1380  
1381 QY TCACTGGACTCAGAGGATGTTGAAAAATAGAGTAAACAGTTTGTCTATGAGAGCCTTCT 1440  
1381 Db TCACTGGACTCAGAGGATGTTGAAAAATAGAGTAAACAGTTTGTCTATGAGAGCCTTCT 1440  
1441 QY GCAACACCTGAGAGTGTGCAAAAGGGTTTCAGATGGAGCAGGACACAAAAAAGTTGAC 1500

1441 Db GCAACACCTGAGAGTGTGCAAAAGGGTTTCAGATGGAGCAGAGACAGAAAAACAAGTTGAC 1500  
1501 QY CGGTCTCTGTGTGGCTCAAAACACTCCGTTCGAGTAGTGTGATGTGAGGCGGATGCATCA 1560  
1501 Db CGGTCTCTGTGTGGCTCAAAACACTCCGTTCGAGTAGTGTGATGTGAGGCGGATGCATCA 1560  
1561 QY GAAAGGCAAGAGGATGGCACCAATGGTGTGAGTGAAGAAACGAATGAAGACACTAATAA 1620  
1561 Db GAAAGGCAAGAGGATGGCACCAATGGTGTGAGTGAAGAAACGAATGAAGACACTAATAA 1620  
1621 QY CCTCAAACTTCAGAGTCCAATGACGCGCAGTAGTGAATCAGCTCCAATATACCCGATCCA 1680  
1621 Db CCTCAAACTTCAGAGTCCAATGACGCGCAGTAGTGAATCAGCTCCAATATACCCGATCCA 1680  
1681 QY TGAAGTCTGTGTCTGACGAGGGTGAATTTGCCCTTCCAAAGCTCTCTTCCAGAGGTA 1740  
1681 Db TGAAGTCTGTGTCTGACGAGGGTGAATTTGCCCTTCCAAAGCTCTCTTCCAGAGGTA 1740  
1741 QY TTGCGCAAAAGTTTACATATCGAGAGAAACACAGAGAGGAAACACACAAACAAAGAA 1800  
1741 Db TTGCGCAAAAGTTTACATATCGAGAGAAACACAGAGAGGAAACACACAAACAAAGAA 1800  
1801 QY CAAAGATATCCAATGGCACTTTGATCTTAACTTTCACAGCTCAGTTAAACACCAAGTTGAT 1860  
1801 Db CAAAGATATCCAATGGCACTTTGATCTTAACTTTCACAGCTCAGTTAAACACCAAGTTGAT 1860  
1861 QY CAAAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAGCTAATG 1920  
1861 Db CAAAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATTTAGATGCTTCAAGCTAATG 1920  
1921 QY AGTAGAGAGAGACAGGTTTTTAAACCATACAAAGATGTTCCATGGAAGCCAAAGAAAGT 1980  
1921 Db AGTAGAGAGAGACAGGTTTTTAAACCATACAAAGATGTTCCATGGAAGCCAAAGAAAGT 1980  
1981 QY AGAATCCTCAACAACTCCTATCATTTCAATGTGGAACAGAAAGATCCCAACCGGATCGG 2040  
1981 Db AGAATCCTCAACAACTCCTATCATTTCAATGTGGAACAGAAAGATCCCAACCGGATCGG 2040  
2041 QY TTGMAACTCAAGCTTCCACATGAGACTCTATTTTTCATCTGATCTGTGTTTGTACTCTG 2100  
2041 Db TTGMAACTCAAGCTTCCACATGAGACTCTATTTTTCATCTGATCTGTGTTTGTACTCTG 2100  
2101 QY TTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTTCAGGCGCTTTGTATTGT 2160  
2101 Db TTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTTCAGGCGCTTTGTATTGT 2160  
2161 QY TTCTTGTCCATAGTCTTCTGTAACTTTGACTCTGTATTATTCAACAAATCATAAACT 2220  
2161 Db TTCTTGTCCATAGTCTTCTGTAACTTTGACTCTGTATTATTCAACAAATCATAAACT 2220  
2221 QY GTTTAATCTTTTTCCTCAAAAAA 2254  
2221 Db GTTTAATCTTTTTCCTCAAAAAA 2254

## RESULT 2

US-08-843-572E-1

; Sequence 1, Application US/08843572E  
; Patent No. 6388172  
; GENERAL INFORMATION:  
; APPLICANT: TOBIN, ELAINE  
; APPLICANT: WANG, ZHI-YONG  
; APPLICANT: SUN, LIN  
; TITLE OF INVENTION: PHYTOCHROME REGULATED  
; TITLE OF INVENTION: TRANSCRIPTION FACTOR FOR CONTROL OF HIGHER PLANT  
; TITLE OF INVENTION: DEVELOPMENT  
; FILE REFERENCE: P601-D-36315.60100  
; CURRENT APPLICATION NUMBER: US/08/843,572E  
; CURRENT FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Microsoft Word 7.0 (DOS text file)  
; SEQ ID NO 1:

; LENGTH: 4344  
 ; TYPE: DNA  
 ; ORGANISM: AR  
 US-08-843-572E-1

Query Match 67.8%; Score 1527.6; DB 4; Length 4344;  
Best Local Similarity 87.9%; Pred. No. 0;  
Matches 1803; Conservative 0; Mismatches 9; Indels 240; Gaps 3;

Qy	441	TTTCTCCAAGTAGAGAAAGGCTGAAGCTAAAGGTGTAGCTATGGTCAAGCGCTAGA	500
Db	2180		
Qy	501	TCTCTCTCAGGTAGAGAAAGGCTGAAGCTTAAAGGTGTAGCTATGGTCAAGCGCTAGA	2239
Db	2240		
Qy	561	CATAGCTATTCTCTCTCCACGGCTAAGCGTAAACCAACAACTCCTTATCCTCGAAAGAC	560
Db	2240		
Qy	561	GGGAAGTGGAAACGATCCTTATGTCAAAAACGGGTGTGAATGATGGAAAAAGAGTCCCTTGG	620
Db	2300	GGGAAGTGGAAACGATCCTTATGTCAAAAACGGGTGTGAATGATGGAAAAAGAGTCCCTTGG	2359
Qy	621	ATCAGAAAAAGTGTGGCATCCG-	643
Db	2360		
Qy	644	-----AGATGGCCAA	650
Db	2420	GTGTCACATTGCTCCTCATGTTATTAATACGATTGTGCTTCGTTTATAGATGGCCAA	2479
Qy	654	TGAAGATCGACAACAATCAAGCCTGAAGAGAAACCTCTGCAGGAAGACAACCTGTTCCAGA	713
Db	2480		
Qy	714	TGAGATCGACAACAATCAAGCCTGAAGAGAAACCTCTGCAGGAAGACAACCTGTTCCAGA	2539
Db	2540		
Qy	774	TTGTTTCACTCATCAGTATCTCTCTGCTGCATCCTCCATGAATAAAAGTTGTATAGAGAC	773
Db	2540		
Qy	774	ATCAAAACGCAAGCACTTTCCGCGAGTCTTTGCGCTTTCACGGGAAG-	817
Db	2600	ATCAAAACGCAAGCACTTTCCGCGAGTCTTTGCGCTTTCACGGGAAGGTAATAAACAAATCT	2659
Qy	818	-----	817
Db	2660	TTCAATTGCTATTGAGGTTTTAAGACGATTAGTACTTTTCATGAACCTAAAAACGCTGGGG	2719
Qy	818	-----AGGGAAGTCAGAATAACAGGGTGAAGAAAGGAGTCAAACCTCAGATTGGAATGCA	870
Db	2720	GAATAACAGGGAAGTCAAGAATAACAGGGTGAAGAAAGGAGTCAAACCTCAGATTGGAATGCA	2779
Qy	871	AAATCTCTGAAAAACGGTAAATGACGAAGGACCTCAGACTTATCCGATGCATATCCCTGTG	930
Db	2780	AAATCTCTGAAAAACGGTAAATGACGAAGGACCTCAGACTTATCCGATGCATATCCCTGTG	2839
Qy	931	CTAGTGCCATTGGGGAGCTCAATAACAAAGTTCTCTATCACATCCTCCTTCAGAGCCAGAT	990
Db	2840	CTAGTGCCATTGGGGAGCTCAATAACAAAGTTCTCTATCACATCCTCCTTCAGAGCCAGAT	2899
Qy	991	AGTCATCCCCACACAGTTGCAGGAGATTATCAGTCGTTTCCTTAATCATATAATGTCAACC	1050
Db	2900	AGTCATCCCCACACAGTTGCAGGAGATTATCAGTCGTTTCCTTAATCATATAATGTCAACC	2959
Qy	1051	CTTTTACAAACACCGGCTCTTTTACTGCGCGAACTTTGCGCTCATCATTTTGGCCTCCC	1110
Db	2960	CTTTTACAAACACCGGCTCTTTTACTGCGCGAACTTTGCGCTCATCATTTTGGCCTCCC	3019
Qy	1111	GATTCTAGTGGTGGCTCACTGTTTCCAGGAACTCACTCCGAATCTGGCTGCCATGGCC	1170
Db	3020	GATTCTAGTGGTGGCTCACTGTTTCCAGGAACTCACTCCGAATCTGGCTGCCATGGCC	3079
Qy	1171	GCAGCCACTGTTGCAGCTGCTAGTGTGTTGGTGGCTGCCAATGATTTACCTTTATGT	1230
Db	3080	GCAGCCACTGTTGCAGCTGCTAGTGTGTTGGTGGCTGCCAATGATTTACCTTTATGT	3139

```
Db      4220 ACCTGGAAGAA 4231      ||| ||
RESULT 3
US-09-202-712-1
; Sequence 1, Application US/09202712
; Patent No. 6265637
; GENERAL INFORMATION:
; APPLICANT: Coupland, George M
; APPLICANT: Schaffer, Robert J
; TITLE OF INVENTION: Genetic control of flowering
; FILE REFERENCE: 620-55
; CURRENT APPLICATION NUMBER: US/09/202,712
; CURRENT FILING DATE: 1999-01-11
; EARLIER APPLICATION NUMBER: PCT/GB97/01676
; EARLIER FILING DATE: 1997-06-23
; EARLIER APPLICATION NUMBER: GB 9613132.1
; EARLIER FILING DATE: 1996-06-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2526
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (338)..(2275)
US-09-202-712-1

Query Match      9.6%; Score 217.2; DB 3; Length 2526;
Best Local Similarity 49.9%; Pred. No. 2.6e-49;
Matches 986; Conservative 0; Mismatches 823; Indels 165; Gaps 11;

Qy      235 GTGATGAGACAAATTCGTCTGGAGAGATCGTGTATTAGACTCGGAAGCCATATACG 294
Db      235 GTTATGATACTAATACATCTGGAGAAGAATATTAGCTAAGGCAAGAAAGCCATATACA 394

Qy      295 ATAAACAAGCAACGTGAAAGGTGGAAGTGAAGAAACATAATAGATTCAATGAAGCTTTG 354
Db      295 ATAAACAAGCAGCAGAGCGATGACTGAGTACGATGAGATGAGAGTTCTTAGAAGCCTTG 454

Qy      355 AGGCTTTATGTTAGAGCATGCGCAAGATTTGAAGAACATGTAGCAACAAACAACTGCTGC 414
Db      455 AGGCTTTATGGAAGAGCTTGGCAACGAATTTGAAGAACATATTGGGCAAAAGAGCTGCTGT 514

Qy      415 CAGATAAGAAGTCACGCTCAGAAATTTTCTCCAAGGTAGAGAAGAGCGTGAAGCTTAA 474
Db      515 CAGATCAGAAGTCATGCACAAAGTTCTTCAAAAGTTGGAGAAGAGGCTGAAGTTAA 574

Qy      475 GGTGTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACCGCCCTAAGCGTAAA 534
Db      575 GGCATCCTGTTTCCCAAGCTTTGGACATAGAAATTCGCGCTCTCTGCTTAAACGAAA 634

Qy      535 CCAACAACATCTTATCTCGAAAGACGGGAAGTGAACGATCTTATGTCAAAAACGGGT 594
Db      635 CCCAATACTCTTATCTCTGAAAACCTTGGGAACACACGATACATCTTCTCTCAAGTATCA 694

Qy      595 GTGAATGATGGAAGAGTCCCTTTGGATCAGAAAAGTGTCCGATCTCGATGCGCCAAT 654
Db      695 TCAGCAAAAGATGCAAAACCTTGTTCATCGGCCCTCTCTTCCACAGTTGGAATCAGCGGTC 754

Qy      655 GAAGATCGACAAACATCAAGCCCTGAAGAGAAAAC-----TCTCAGGAA 699
Db      755 TTGATTGGAAGAAAATGCCGTTCTCTGAGAAAACATCAACTGGAAAAGAAATCAAGAT 814

Qy      700 GACAACTGTTTCAGATTGTTTCACTCATCATGATCTCTCTGCTGATCTCTCATGAAATAA 759
Db      815 GAGAAATTCGCGGTGTTTCTACTGTGAACAAGTATCCCTTATCAACGAAACAGGTAAGT 874

Qy      760 AGTTGTATAGACATCAAAACGCAAGCACTTT-----CCGCGAGTCTTCGCTTC 809
Db      875 GCCGACATTGAAACAGTAAAGACCTCACTGTGTGACAAACCGGGTTCAAGATGTTCCCAAG 934

Qy      810 ACGGGAAGCGGAAGTCAGAAATACAGGGTA-----AGAAAGAGTCAAAAC 855
Db      935 AAGAACAAAGACAAAGATGGTAACGATGGTACTACTGTGCACAGCATCAAAACATACCT 994

Qy      856 TCAGATTGTAATGCAAAATCTCTGGAACCGGTAAATGAGCAAGGACCTCAGACTTTATCC- 914
Db      995 TGGCATTTCCACGCGAGATATTGTGAACGGGAATATAGCAAAATGCCCTCAAAATCATCCC 1054

Qy      915 -----G 915

Db      1055 TCAGGTATGTTATCTCAAGACTTTCATGTTTCATCTATGAGAGAAGAAACTCACGGCAC 1114

Qy      916 ATGCATATCCCTGTGCTAGTGCCTATGGGAGCTCAATAACAAAGTCTCTATCATCATCT 975
Db      1115 GCAAAATCTTCAAGCTACAACAGCATCTGCTACTACTACAGCTTCTCATCAAGCGTTTCCA 1174

Qy      976 CTTTCAGAGCCAGATAGTCATCCCCACACAGCTTGCAGGAGATTATCAGTCGTTTCTCTAAT 1035
Db      1175 GCTTGTCAATCACAGGATGATTACGGTTCGTTTCTCCAGATATCATCTACTTCTCCAAAT 1234

Qy      1036 CATATAATGCAACCCCTTTTACAACACCGGCTCTTTATATCTGCGGCAACTTTCGCCTCA 1095
Db      1235 CTTATTATGCAACTCTCTACAGAAATCTGCAGCTCATGCTGCAGCTACATTCGCTGCT 1294

Qy      1096 TCATTTTGGCTCCCG-----ATTCTAGTGGTGGCTCACCTGTTCCAGGGGAAC 1143
Db      1295 TCGGTCTGGCCCTTATGCGAGTGTGCGGAATTTCTGGTGATTCATCAACCCCAATGAGTCT 1354

Qy      1144 TCACCTCCGAATCTGGCTGCGCATGGCCGACGCACTGTTGCAGCTGCTAGTGTCTGGTGG 1203
Db      1355 TCTCCTCCAAGTATAACTGCCATTTGCCGCTGCTACAGTAGCTGTGCAACTGCTTGGTGG 1414

Qy      1204 GCTGCCAATGGAATTAATACCTTTATGCTCCTCTTATGTTAGTTCAGGTGGTTTCACTAGTCA 1263
Db      1415 GCTTCTCATGAGACTTCTTCTGTATGGCTCCAGTCCCAATAACATGTTGTTCCATTCTCA 1474

Qy      1264 CTTCCATCTACTTTTGGACCATCATGTGATGATGATACACAAAGCAAGCACTTTACNA 1323
Db      1475 ACTGTTGAGTTTCCAATCTCCAGCAATGACTGAATGGATACGTTTGAATAACTCAACCG 1534

Qy      1324 CATGGTTCTGTGCAGAGCCGAGACAAACACCTCCGAGGCACTCAAAGGCTCGATCTTCA 1383
Db      1535 TTTGAGAAACAAACACACAGCTCTCGAAGATCAAACTTGGCTTCGAATCTCCAGCTTCA 1594

Qy      1384 CTGACTCAGAGGATGTTGAAAAT-----AAGATAAACAGTTCATGAGGAG 1434
Db      1595 TCATCTGATGATTTCAGATGAGACTGGAGTAAACCAAGCTAAATGCCGACTCAAAAACCAAT 1654

Qy      1435 CCTTCTGCAACACCTGAGAGTGATGCAA-----AGGGTTTCAGATGGAGCAGGA 1482
Db      1655 GATGATAAAATTCAGGAGGTTGTTGTTACTGCCCCGCTGTGCATGACTCAAACTGCCCG 1714

Qy      1483 GACAGAAAAACAAGTTGACCGGCTCCTCGTGTGGCTCAAAACACTCCGTCGAGTAGTGATGAT 1542
Db      1715 AAGAAAAATCTTGTGGACCGCTCATCGTGGCTCAATAACACTTTCAGGAGGTGACCGCA 1774

Qy      1543 GTTGAGCGGATGATCAGAAAAGCAAGAGATGGCAACCAATGTGAGGTGAAAGAAACG 1602
Db      1775 GAAACTGATGATTTAGATAAAATGGAGAAGATAAAGAGGATGTGAAGGAGACAGATGAG 1834

Qy      1603 AAT-----GAAGACACTTAATAACCTCAAACTTCAGAGTCCAATGCACGCGCGAGTAGA 1656
Db      1835 AATCAGCCAGATGTTTATGAGTTAAATAACCGTAAGATTAAATGAGAGACAACACAGC 1894

Qy      1657 ATCAGCTCCAATATAACCGGCTCCATGGAAGTCTGTCTGACGAGGGTTCGAATTTGCCTTC 1716
Db      1895 AACAACAATGCAACTACTGATTCTGTGGAAGGAAGTCTCCGAGAGGGTTCGTATACGGTTT 1954

Qy      1717 CAAGCTCTCTTCCAGAGAGGATTATCCCGAAAGTTTACATATCAGAGAGAAACACAGA 1776
Db      1955 CAGGCTCTCTTTGCAAGAGAAAGATTTCCTCAAAAGCTTTTCGCTCTCTCAAGTGGCAGAG 2014
```





NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELEPHONE: (703)836-9300  
TELEFAX: (703)836-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZgpt-F1s  
US-08-232-463-14

Query Match 2.6%; Score 59.2; DB 1; Length 7218;  
Best Local Similarity 3.4%; Pred. No. 6.3e-06;  
Matches 13; Conservative 225; Mismatches 148; Indels 0; Gaps 0;  
QY 186 AAAGTGAATTTGTAAGAGAGATCTGAAGTTGTAGAGGAGCTTAGTGATGAGAC 245  
Db 1451 ATAGAAGAATTTGTAACRTRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1392  
QY 246 AATTCCTCGGAGAGATCTGGTTATTAAGACTCGGAAGCCATATACGATAACAAGCA 305  
Db 1391 RRR 1332  
QY 306 ACGTGAAGTGAAGTGAAGAGAACATAAGATTGAGCTTTGAGGCTTTATGG 365  
Db 1331 RRR 1272  
QY 366 TAGAGATCGCAGAGATTAAGAACATGTAGCAACAACTGCTGCCAGATAAGAAG 425  
Db 1271 RRR 1212  
QY 426 TCACGCTCAAGAAATTTTCTCAAGTAGAGAGAGCTGAAGCTAAAGGTGATGCTAT 485  
Db 1211 RRR 1152  
QY 486 GGTCAAGCGCTAGACATAGTATCTCTCCAGCGCTAAGCTAAGCTAACAACATCC 545  
Db 1151 RRR 1092  
QY 546 TTATCTCGAAAGACGGGAAGTGGAA 571  
Db 1091 RRR 1066

RESULT 7  
US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELEPHONE: (703)836-9300  
TELEFAX: (703)836-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZgpt-F1s  
US-08-232-463-14  
Query Match 2.2%; Score 49.8; DB 1; Length 7218;  
Best Local Similarity 1.6%; Pred. No. 0.0023;  
Matches 6; Conservative 221; Mismatches 148; Indels 0; Gaps 0;  
QY 910 TATCCGATCATATCCTGTCTAGTGGCATTGGGAGCTCAATAACAAGTTCTCTATCA 969  
Db 1066 TTYY 1125  
QY 970 CATCTCTTCAGAGCAGATAGTATCCCAACAGTTGCAGGAGATTATCAGTCGTT 1029  
Db 1126 YY 1185  
QY 1030 CCTAATCATATGATCAACCTTTTACAAACACCGCTCTTTATCTACTCGCGCACTTC 1089  
Db 1186 YY 1245  
QY 1090 GCCTCATCATTTGGCTCCGCTCCTAGTGGTGGCTCCTGTTCCAGGAACCTCACCT 1149  
Db 1246 YY 1305  
QY 1150 CCGAATCTGGCTGCCATGCCCGCAGCTGTTGCAGCTGCTAGTCTGTTGGTGGCTGCC 1209  
Db 1306 YY 1365  
QY 1210 AATGATATTACCTTTATGCTCTCTAGTTCAGTGGTGGTTCAGTCTATCTCCTCA 1269  
Db 1366 YY 1425  
QY 1270 TCTACTTTTGACCA 1284  
Db 1426 YYYYYYYYGTACCA 1440

RESULT 8  
5273901-6/c  
Patent No. 5273901  
APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,  
SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;  
AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.  
TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS  
SPOROZOITE 21.5 KB ANTIGEN, AC-6B  
NUMBER OF SEQUENCES: 11  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/581,693  
FILING DATE: 12-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 215,162

Query Match	1.9%;	Score 42.6;	DB 1;	Length 1974;
Best Local Similarity	51.3%;	Pred. No. 0.11;		
Matches 99;	Conservative 0;	Mismatches 94;	Indels 0;	Gaps 0;
2059	ACATGAGAGCTCTATT	TTTCATCTGATCTCT	CTGTTGTACTCTGTTT	TAAAGTTTCAAGAC 2118
b	1317	AGATGCTCTTTGGAT	TCGATGCTCTTTTCCAT	CTCTTTTAAATGCATGTAA 1258
2119	CACCTGCTACATTTT	CTCTTTTCTTTTGAGG	CCCTTTGTATTTGT	TGTTCTCTTGCCATPAGTCTT 2178
y				
b	1257	TCTTGTGTTCTTTT	CTTTTCTTTTCTTTT	TTTTTTTTTTTTTTTCTAAGATTGCTCT 1198
2179	CTGTAAACATTTGAC	TCTGTATTAATCAACA	AAATCATAACTGTTT	TAATCTTTTTTTTTTTC 2238
y				
b	1197	ACTTTTAGTAATATT	TGTAATATATACCT	TCTATTTCATTAATGATTTACAAG 1138
2239	CAAAAAAAAAAAAA	2251		
y				
b	1137	TTAGTAAAAAAAAA	1125	

RESULT 10  
S-08-971-395-3/C  
Sequence 3, Application US/08971395  
Patent No. 6359197

TITLE OF INVENTION: Transgenic Plants with Altered  
 Senescence Characteristics  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Quarles & Brady  
 STREET: 1 South Pinckney Street  
 City: Madison  
 STATE: WI  
 COUNTRY: US  
 ZIP: 53701-2113  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/971.395

	Query Match	1.9%	Score 42.6	DB 4	Length 1974
	Best Local Similarity	51.3%	Pred. No. 0.11		
	Matches 99	Conservative 0	Mismatches 94	Indels 0	Gaps 0
2059	ACATGAGACGCTATTTTCATCTGATCTGTTGTTGACTCTGTTTTAAAGTTTTC	CAAGAC	2118		
1317	AGATGCTCTCTTGGGATTCGATAGCTCGGATCTCTTCATCTCTTTTAAATGCATTGTAA	1258			
2119	CACGTGCTACATTTTCTTTTTCTTTTGGGCGCTTTGTATTGTTTCCTCTGCATAGTCTT	2178			



```

Db      814  CTATATATTTTATTTCTCAAAATATACTAAAA  848

RESULT 13
US-09-253-691-3/c
; Sequence 3, Application US/09253691
; Patent No. 6124100
; GENERAL INFORMATION:
; APPLICANT: Dong Kyu JIN
; TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
; TITLE OF INVENTION: Using Trinucleotide Repeats Sequence
; FILE REFERENCE: 1942/36
; CURRENT APPLICATION NUMBER: US/09/253.691
; CURRENT FILING DATE: 1999-02-22
; EARLIER APPLICATION NUMBER: KR 98-6,278
; EARLIER FILING DATE: 1996-02-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: WordPerfect 6.1/Windows
; SEQ ID NO 3
; LENGTH: 397
; TYPE: DNA
; ORGANISM: human
US-09-253-691-3

```

Query Match	1.8%	Score 41;	DB 3;	Length 397;
Best Local Similarity	47.5%;	Pred. No. 0.14;		
Matches 122;	Conservative 0;	Mismatches 135;	Indels 0;	Gaps 0;
Qy	1023	GTCCGTTCCCTAAATCATATATGTCAACCCCTTTTAAACACCGGCTCTTTATCTACTGCCGC	1082	
Db	396	GGCCTTTTCACATGGATGAACTCTGTGCTGATAGGTCCTCCGCTGCTGCTGCTGCTGCTGC	337	
Qy	1083	AACCTTCGCCTCATCTATTTTGGCCTCCGATCTTAGTGGTGCTCACTGTTCAGGGAA	1142	
Db	336	TGC	277	
Qy	1143	CTACCTCCGAATCTGGCTGCATGGCGCGACGACCTGTTGCACCTGCTAGTGTCTGGTG	1202	
Db	276	TGC	217	
Qy	1203	GGCTGCCAATGGATTATACCTTTATGTGCTCCTCTTAGTTCAGGTGGTTTCACTAGTCA	1262	
Db	216	TGC	157	
Qy	1263	TCCTCCCATCTACTTTTG	1279	
Db	156	TGCTGCTGCTGCTGCTG	140	

RESULT 14  
US-09-014-969-14/c  
; Sequence 14, Application US/09014969  
; Patent No. 5965397  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: LaVallie, Edward R.  
; APPLICANT: Racie, Lisa A.  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Spaulding, Vikki  
; APPLICANT: Agostino, Michael J.  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; TITLE OF INVENTION: ENCODING THEM  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/014,969
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-014-969-14

```

Query Match	1.8%	Score 41;	DB 2;	Length 2447;
Best Local Similarity	52.0%	Pred. No. 0.34;		
Matches 92;	Conservative 0;	Mismatches 85;	Indels 0;	Gaps 0;
Qy 2068	TCATATTTTCATCTGATCTGGTTGTTGACTCTCTGTTTTTAAGTTTTTCAAGACCACCTGCTAC	2127		
Db 2412	TT	2353		
Qy 2128	ATTTCTTTTTTCTTTTTTGAGGCCCTTGTATTGTTTCTCTTGCCATAGTCTCTCTGTAACA	2187		
Db 2352	TT	2293		
Qy 2188	TTTGACTCTGTATTATTCACAAATCATAAACTGTTTAAATCTTTTTTTTTTCCAAAAA	2244		
Db 2292	TT	2236		

```

RESULT 15
US-09-181-585-3
; Sequence 3, Application US/09181585
; Patent No. 6524791
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Koob, Michael
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 11000900101
; CURRENT APPLICATION NUMBER: US/09/181,585
; CURRENT FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA
; OTHER INFORMATION: comprising exons E, C, and A
US-09-181-585-3

```

	Query Match	1.8%	Score 40.8	DB 4	Length 1037
	Best Local Similarity	47.0%	Pred. No. 0.25		
	Matches 126	Conservative 0	Mismatches 142	Indels 0	Gaps 0
Qy	1027	TTTCCTAATCATATAATGTCACCCCTTTTACAAACACCGGCTCTTTATATCTGCCGAAC	1086		
Db	634	TTTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTCTGCTGCTGCTGCT	693		
Qy	1087	TTGCCTCATCATTTTGGCGCTCCCGATTCCTAGTGGGCTACCTGTTCCAGGAAC	1146		
Db	694	GTGCT	753		



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2254	100.0	2254	14	US-10-084-553-3	Sequence 3, Appli
2	2240	99.4	2240	10	US-09-887-576-617	Sequence 617, App
3	1827	81.1	1827	10	US-09-938-842A-1500	Sequence 1500, Ap
4	1527.6	67.8	4344	14	US-10-084-553-1	Sequence 1, Appli
5	479	21.3	479	11	US-09-770-961-961	Sequence 961, App
6	217.2	9.6	2526	11	US-09-934-455-119	Sequence 119, App
7	217.2	9.6	2526	15	US-10-278-536-199	Sequence 199, App
8	213	9.4	1938	8	US-09-938-842A-1116	Sequence 1116, Ap
9	189	8.4	1134	10	US-09-938-842A-4192	Sequence 4192, Ap
10	189	8.4	1137	10	US-09-887-576-702	Sequence 702, App
11	116.8	5.2	1164	11	US-09-934-455-247	Sequence 247, App
12	83	3.7	1237	14	US-10-395-403-49	Sequence 49, Appl
13	74.8	3.3	365	10	US-09-878-574-707	Sequence 707, App
14	55.2	2.4	664	12	US-09-814-353-4739	Sequence 4739, Ap
15	55.2	2.4	664	12	US-09-814-353-11036	Sequence 11036, A
16	53	2.4	397	10	US-09-960-352-113784	Sequence 113784, A

QY 121 GAATTAATAATGGAATCTTTATCGAATCCAAAGCTGATTTTGTCTTTTCATTTGAATCATC 180  
DB 121 GAATTAATAATGGAATCTTTATCGAATCCAAAGCTGATTTTGTCTTTTCATTTGAATCATC 180  
QY 181 TCTCTAAAGTGGAAATTTTGTAAAGAGAGATCTGAAGTGTGTAGAGAGCTTAGTGATG 240  
DB 181 TCTCTAAAGTGGAAATTTTGTAAAGAGAGATCTGAAGTGTGTAGAGAGCTTAGTGATG 240  
QY 241 GAGACAAATTCGCTCGAGAGATCTGGTTATTAAAGACTCGGAGCCATATACGATTAACA 300  
DB 241 GAGACAAATTCGCTCGAGAGATCTGGTTATTAAAGACTCGGAGCCATATACGATTAACA 300  
QY 301 AAGCAACGTGAAAGGTGGACTGAGCAAGAAACATAATAGATTTCAATTGAAGCTTTGAGGCTT 360  
DB 301 AAGCAACGTGAAAGGTGGACTGAGCAAGAAACATAATAGATTTCAATTGAAGCTTTGAGGCTT 360  
QY 361 TATGGTAGAGCATGGCAGAGATTTGAAGAACATGTAGCAACAAAACTGCTGTCAGATA 420  
DB 361 TATGGTAGAGCATGGCAGAGATTTGAAGAACATGTAGCAACAAAACTGCTGTCAGATA 420  
QY 421 AGAAGTCACCTCAGAAATTTTCTCCAAGGTAGAGAAAGGCTGAAGCTTAAAGGTGA 480  
DB 421 AGAAGTCACCTCAGAAATTTTCTCCAAGGTAGAGAAAGGCTGAAGCTTAAAGGTGA 480  
QY 481 GCTATGGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCTTAAGCGTAAACCAAAC 540  
DB 481 GCTATGGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCTTAAGCGTAAACCAAAC 540  
QY 541 AATCCTTATCTCGAAAGAGCGGAAGTGGAAACGATCTTATGTCAAAAAACGGGTGTGAAT 600  
DB 541 AATCCTTATCTCGAAAGAGCGGAAGTGGAAACGATCTTATGTCAAAAAACGGGTGTGAAT 600  
QY 601 GATGGAAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCGAGATGGCCAAATGAAGT 660  
DB 601 GATGGAAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCGAGATGGCCAAATGAAGT 660  
QY 661 CGACAACAATCAAGCGCTGAAGAGAAAACCTCTGCAGGAAGACAACCTGTTTCAAGATTGTTTC 720  
DB 661 CGACAACAATCAAGCGCTGAAGAGAAAACCTCTGCAGGAAGACAACCTGTTTCAAGATTGTTTC 720  
QY 721 ACTCATCAGTATCTCTGCTGATCTCCATGAATAAAAGTGTATAGAGACATCAAAC 780  
DB 721 ACTCATCAGTATCTCTGCTGATCTCCATGAATAAAAGTGTATAGAGACATCAAAC 780  
QY 781 GCAAGCACTTTCGCGAGTTCTTGCCCTTACGGAAGAGGGAAGTCAAGATAACAGGTA 840  
DB 781 GCAAGCACTTTCGCGAGTTCTTGCCCTTACGGAAGAGGGAAGTCAAGATAACAGGTA 840  
QY 841 AGAAAGAGTCAAACTCAGATTTGAATGCAAAATCTCTGAAAAACGGTAAAGCAAGGA 900  
DB 841 AGAAAGAGTCAAACTCAGATTTGAATGCAAAATCTCTGAAAAACGGTAAAGCAAGGA 900  
QY 901 CCTCAGACTTATCCGATGATATCCCTGCTAGTGCCATTTGGGGAGCTCAATAACAAGT 960  
DB 901 CCTCAGACTTATCCGATGATATCCCTGCTAGTGCCATTTGGGGAGCTCAATAACAAGT 960  
QY 961 TCTCTATCAGATCTCTTTCAGAGCCAGATAGTCAATCCCAACAGTTGAGGAGATTAT 1020  
DB 961 TCTCTATCAGATCTCTTTCAGAGCCAGATAGTCAATCCCAACAGTTGAGGAGATTAT 1020  
QY 1021 CAGTCGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTACTGCC 1080  
DB 1021 CAGTCGTTTCTTAATCATATAATGTCAACCCCTTTTACAAACACCGGCTCTTTTACTGCC 1080  
QY 1081 GCAACTTTCGCTCATCATTTTGGCTCCCGATTTCTAGTGGTCCCTCACCTGTTCCAGGG 1140  
DB 1081 GCAACTTTCGCTCATCATTTTGGCTCCCGATTTCTAGTGGTCCCTCACCTGTTCCAGGG 1140  
QY 1141 AACTCACTCCGAATCTGGCTGCATGGCGGAGCCACTGTTGACGCTGCTAGTCTGG 1200  
DB 1141 AACTCACTCCGAATCTGGCTGCATGGCGGAGCCACTGTTGACGCTGCTAGTCTGG 1200

QY 1201 TGGCTGCCAATGGATTATTACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGT 1260  
DB 1201 TGGCTGCCAATGGATTATTACCTTTATGTGCTCCTCTTAGTTTCAGGTGGTTTCACTAGT 1260  
QY 1261 CATCTCCATCTACTTTTGGACCATCATGTGTAGTAGTACACAAAGCAAGCACTTTA 1320  
DB 1261 CATCTCCATCTACTTTTGGACCATCATGTGTAGTAGTACACAAAGCAAGCACTTTA 1320  
QY 1321 CAACATGGTTCTGTGAGAGCCGAGACCAAGAACACTCCGAGGCATCAAAGGCTCGATCT 1380  
DB 1321 CAACATGGTTCTGTGAGAGCCGAGACCAAGAACACTCCGAGGCATCAAAGGCTCGATCT 1380  
QY 1381 TCACTGGACTCAGAGGATTTGAAAAAAGAGTAAACACAGTTTCTCATGACAGCCTTCT 1440  
DB 1381 TCACTGGACTCAGAGGATTTGAAAAAAGAGTAAACACAGTTTCTCATGACAGCCTTCT 1440  
QY 1441 GCAACACTCGAGAGTGTCAAAAGGGTTTCAGATGGAGCAGAGACAGAAAAAAGTTGAC 1500  
DB 1441 GCAACACTCGAGAGTGTCAAAAGGGTTTCAGATGGAGCAGAGACAGAAAAAAGTTGAC 1500  
QY 1501 CGGTCTCGTGTGGCTCAAAACACTCCGTCAGTAGTGTATGTTGAGGCGGATGCATCA 1560  
DB 1501 CGGTCTCGTGTGGCTCAAAACACTCCGTCAGTAGTGTATGTTGAGGCGGATGCATCA 1560  
QY 1561 GAAAGGCAAGAGGATGGCACCAATGGTGAAGTGAAGAAAAAGTGAAGACACTAAATAA 1620  
DB 1561 GAAAGGCAAGAGGATGGCACCAATGGTGAAGTGAAGAAAAAGTGAAGACACTAAATAA 1620  
QY 1621 CCTCAAACTTCAGAGTCCAAATGCAACGCGCAGTAGAATCAGTCCCAATATAAACCGATCCA 1680  
DB 1621 CCTCAAACTTCAGAGTCCAAATGCAACGCGCAGTAGAATCAGTCCCAATATAAACCGATCCA 1680  
QY 1681 TGAAGTCTGTGTCTGACGAGGCTCGAATTTGCCCTTCCAAGCTCTCTTCTCCAGAGAGTA 1740  
DB 1681 TGAAGTCTGTGTCTGACGAGGCTCGAATTTGCCCTTCCAAGCTCTCTTCTCCAGAGAGTA 1740  
QY 1741 TTCCCGCAAAAGTTTACATATCGAAGAAACACAGAGAGGAAGAAACAAACAACAAGAA 1800  
DB 1741 TTCCCGCAAAAGTTTACATATCGAAGAAACACAGAGAGGAAGAAACAAACAACAAGAA 1800  
QY 1801 CAAAGATATCCAATGGCACTTGATCTTAACTTCAAGCTCAGTTAAACACCGATTAATG 1860  
DB 1801 CAAAGATATCCAATGGCACTTGATCTTAACTTCAAGCTCAGTTAAACACCGATTAATG 1860  
QY 1861 CRAAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATAGATGCTTCAAAGCTAATG 1920  
DB 1861 CRAAGAGAGAGAGAAACACAGGATTTCTTGGAAATCGGATAGATGCTTCAAAGCTAATG 1920  
QY 1921 AGTAGAGAGAGAAACAGGTTTAAACCATACAAAGATGTTTCCATGGAAGCCAAAGAAAGT 1980  
DB 1921 AGTAGAGAGAGAAACAGGTTTAAACCATACAAAGATGTTTCCATGGAAGCCAAAGAAAGT 1980  
QY 1981 AGAATCCTCAACAAACATCTATCATTTCAATGTGGAACAGAAAGATCCCAACCGATCGG 2040  
DB 1981 AGAATCCTCAACAAACATCTATCATTTCAATGTGGAACAGAAAGATCCCAACCGATCGG 2040  
QY 2041 TTGGAAACTCAAGCTTCCACATGAGACTCTATTTTTCATCTGATCTGTTTGTGACTCTG 2100  
DB 2041 TTGGAAACTCAAGCTTCCACATGAGACTCTATTTTTCATCTGATCTGTTTGTGACTCTG 2100  
QY 2101 TTTTAAAGTTTTCAGAACCACTGCTACATTTTCTTTTTCAGGCGCTTTGTTGTTTGT 2160  
DB 2101 TTTTAAAGTTTTCAGAACCACTGCTACATTTTCTTTTTCAGGCGCTTTGTTGTTTGT 2160  
QY 2161 TTTCTTGTCCATAGTCTTCTGTAACATTTGACTCTGTATTTTCAACAAATCATAACT 2220  
DB 2161 TTTCTTGTCCATAGTCTTCTGTAACATTTGACTCTGTATTTTCAACAAATCATAACT 2220  
QY 2221 GTTTAATCTTTTTCAGAAAAAAGAAAAA 2254  
DB 2221 GTTTAATCTTTTTCAGAAAAAAGAAAAA 2254





QY 1681 TGAAGTCTGTGCTGACGAGGTCGAATTCGCTTCAAGCTCTTCTCCAGAGGTA 1740  
DB 1681 TGAAGTCTGTGCTGACGAGGTCGAATTCGCTTCAAGCTCTTCTCCAGAGGTA 1740  
QY 1741 TTCCGCAAGTCTTACATATCAGAGAGAACACAGAGAGAACACACACACAGAA 1800  
DB 1741 TTCCGCAAGTCTTACATATCAGAGAGAACACAGAGAGAACACACACACAGAA 1800  
QY 1801 CAAAGATATCAATGGCACTTGATCTTAACTTTCACAGCTCAGTTAACACAGTTGATGAT 1860  
DB 1801 CAAAGATATCAATGGCACTTGATCTTAACTTTCACAGCTCAGTTAACACAGTTGATGAT 1860  
QY 1861 CAAGAGAGAGAGAGAACACAGAGATTTCTTGGAACTCGGATAGATGCTTCAAGCTTAATG 1920  
DB 1861 CAAGAGAGAGAGAGAACACAGAGATTTCTTGGAACTCGGATAGATGCTTCAAGCTTAATG 1920  
QY 1921 AGTAGAGAGAGAGAGTCTTAAACCATACAAAGATGTTCCATGGAGCCAAAGAAAGT 1980  
DB 1921 AGTAGAGAGAGAGAGTCTTAAACCATACAAAGATGTTCCATGGAGCCAAAGAAAGT 1980  
QY 1981 AGAATCTCAACAACTATCATTCATGTTGGAACAGAGAGATCCCAACCGATCGCG 2040  
DB 1981 AGAATCTCAACAACTATCATTCATGTTGGAACAGAGAGATCCCAACCGATCGCG 2040  
QY 2041 TTGGAACCTCAAGCTTCCACATGAGCTCTATTTTCATCTGATCTGTTGTTGATCTCTG 2100  
DB 2041 TTGGAACCTCAAGCTTCCACATGAGCTCTATTTTCATCTGATCTGTTGTTGATCTCTG 2100  
QY 2101 TTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTGGGCTTTCTGATTTGT 2160  
DB 2101 TTTTAAAGTTTCAAGACCACTGCTACATTTCTTTTCTTTTGGGCTTTCTGATTTGT 2160  
QY 2161 TTCTTGTGCATAGTCTTCTGTAACATTTGACTCTGTATTTTCAACAACTATAAAT 2220  
DB 2161 TTCTTGTGCATAGTCTTCTGTAACATTTGACTCTGTATTTTCAACAACTATAAAT 2220  
QY 2221 GTTAAATCTTTTCTTCCA 2240  
DB 2221 GTTAAATCTTTTCTTCCA 2240

## RESULT 3

US-09-938-842A-1500  
; Sequence 1500, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1500  
; LENGTH: 1827  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1500

Query Match 81.1%; Score 1827; DB 10; Length 1827;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 238 ATGGAGACAAATTCGTCTGGAGAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA 297

DB 1 ATGGAGACAAATTCGTCTGGAGAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA 60  
QY 298 ACAAGCAAGCTGAAAGGTGGACTGAGGAAGAACATAATAGATTCTTGAAGCTTTGAGG 357  
DB 61 ACAAGCAAGCTGAAAGGTGGACTGAGGAAGAACATAATAGATTCTTGAAGCTTTGAGG 120  
QY 358 CTTTATGGTAGACATCGGCAGAAAGATTGGAAGAACATGTAGCAACAAAACTGCTGTCCAG 417  
DB 121 CTTTATGGTAGACATCGGCAGAAAGATTGGAAGAACATGTAGCAACAAAACTGCTGTCCAG 180  
QY 418 ATAAGAGTCACCGCTCAGAAATTTTCTCAAGGTAGAGAAAGAGGCTGAAGCTTAAAGT 477  
DB 181 ATAAGAGTCACCGCTCAGAAATTTTCTCAAGGTAGAGAAAGAGGCTGAAGCTTAAAGT 240  
QY 478 GTAGCTATGGGTCAAGGCTAGACATAGCTTATTCCTCTCCACGGCTTAAGCTTAAACCA 537  
DB 241 GTAGCTATGGGTCAAGGCTAGACATAGCTTATTCCTCTCCACGGCTTAAGCTTAAACCA 300  
QY 538 AACCAATCCTTATCCTCGAAAGACGGGAAGTGGAAACGATCCTTATGTCAAAAACGGGTGTG 597  
DB 301 AACCAATCCTTATCCTCGAAAGACGGGAAGTGGAAACGATCCTTATGTCAAAAACGGGTGTG 360  
QY 598 AATGATGGAAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGATGGCCAAATGAA 657  
DB 361 AATGATGGAAAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCCTGAGATGGCCAAATGAA 420  
QY 658 GATCGAACAACTCAAGGCTGAAGAAAACTCTCGCAGGAAGACAACTGTTTCAGATTGT 717  
DB 421 GATCGAACAACTCAAGGCTGAAGAAAACTCTCGCAGGAAGACAACTGTTTCAGATTGT 480  
QY 718 TTCACCTCATCAGTATCTCTGCTGATCCTCCATGAATAAAAGTTGTATAGAGACATCA 777  
DB 481 TTCACCTCATCAGTATCTCTGCTGATCCTCCATGAATAAAAGTTGTATAGAGACATCA 540  
QY 778 AACCGAAGCACTTTCGCGGAGTTCTTGCCTTACGGGAAGAGGGAAGTCAGATAAACAGG 837  
DB 541 AACCGAAGCACTTTCGCGGAGTTCTTGCCTTACGGGAAGAGGGAAGTCAGATAAACAGG 600  
QY 838 GTAAGAAAGAGTCAAACTCAGATTTTGAATGCAAAATCTCTGGAACCGGTATGAGCAA 897  
DB 601 GTAAGAAAGAGTCAAACTCAGATTTTGAATGCAAAATCTCTGGAACCGGTATGAGCAA 660  
QY 898 GGACCTCAGACTTATCCGATGCATATCCCTGCTAGTGCCATTTGGGAGCTCAATAACA 957  
DB 661 GGACCTCAGACTTATCCGATGCATATCCCTGCTAGTGCCATTTGGGAGCTCAATAACA 720  
QY 958 AGTTCTCTATCATCCTCTTCAGAGCCAGATAGTCATCCCAACAGTTGAGGAGAT 1017  
DB 721 AGTTCTCTATCATCCTCTTCAGAGCCAGATAGTCATCCCAACAGTTGAGGAGAT 780  
QY 1018 TATCAGTCTGTTTCTTAATATATATGTCACCCCTTTTACAAACACCGGCTCTTTTACT 1077  
DB 781 TATCAGTCTGTTTCTTAATATATATGTCACCCCTTTTACAAACACCGGCTCTTTTACT 840  
QY 1078 GCGCAACTTTCGCTCTCATCTTTGGCTTCCGATTTAGTGGTGGCTCACCTGTGTCCA 1137  
DB 841 GCGCAACTTTCGCTCTCATCTTTGGCTTCCGATTTAGTGGTGGCTCACCTGTGTCCA 900  
QY 1138 GGAAGTCACTCCGAACTGCTGCTGCGGAGCCAGCTGTTGAGCTGCTAGTGTCT 1197  
DB 901 GGAAGTCACTCCGAACTGCTGCTGCGGAGCCAGCTGTTGAGCTGCTAGTGTCT 960  
QY 1198 TGGTGGGCTGCAATGGATTATACCTTTATGCTCTCTTAGTTTTCAGTGGTGTCACT 1257  
DB 961 TGGTGGGCTGCAATGGATTATACCTTTATGCTCTCTTAGTTTTCAGTGGTGTCACT 1020  
QY 1258 AGTCATCTCTCACTTACTTTTGGACCATCATGTGTAGTAGAGTACACAAAGCAACACT 1317  
DB 1021 AGTCATCTCTCACTTACTTTTGGACCATCATGTGTAGTAGAGTACACAAAGCAACACT 1080  
QY 1318 TTACAACTGTTTGTGTCAGAGCCGAGCAAGAACACTCCGAGGCTCAAGGCTCGA 1377

Db 1081 TTACAAACATGGTCTGTGTCAGAGCCGAGAGCAAGAACCACTCCGAGGCATCAAGGCTCGA 1140  
Qy 1378 TCCTTCACTGCACTCAGAGGATGTTGAAAAATAAGAGTAACCCAGTTTGTCTATGACGACCT 1437  
Db 1141 TCCTTCACTGCACTCAGAGGATGTTGAAAAATAAGAGTAACCCAGTTTGTCTATGACGACCT 1200  
Qy 1438 TCTGCAACACCTCAGAGTGTATGCAAGGGTTTCAGATGGAGCAGAGACAGAAAAACAAGTT 1497  
Db 1201 TCTGCAACACCTCAGAGTGTATGCAAGGGTTTCAGATGGAGCAGAGACAGAAAAACAAGTT 1260  
Qy 1498 GACCGTCTCTGTGCTCAACACCTCCGTCGAGTAGTGTATGATGTTGAGGCGGATGCA 1557  
Db 1261 GACCGTCTCTGTGCTCAACACCTCCGTCGAGTAGTGTATGATGTTGAGGCGGATGCA 1320  
Qy 1558 TCAGAAAGGCAAGAGGATGGCCCAATGGTGGAGTGAAGAAACCAATGAAGACACTAAT 1617  
Db 1321 TCAGAAAGGCAAGAGGATGGCCCAATGGTGGAGTGAAGAAACCAATGAAGACACTAAT 1380  
Qy 1618 AAACCTCAAACTTCAGAGTCAATGCAAGCGGCGAGTAGAATCAGCTCCAATATAACCGAT 1677  
Db 1381 AAACCTCAAACTTCAGAGTCAATGCAAGCGGCGAGTAGAATCAGCTCCAATATAACCGAT 1440  
Qy 1678 CCATGGAAGTCTGTCTGTCGAGGCTGCAATGCGCTTCCAAAGCTCTTCTCCAGAGAG 1737  
Db 1441 CCATGGAAGTCTGTCTGTCGAGGCTGCAATGCGCTTCCAAAGCTCTTCTCCAGAGAG 1500  
Qy 1738 GTATTGCGCAAAAGTTTTACATATCGAGAGAACACACAGAGGAGAAACCAACAAACAA 1797  
Db 1501 GTATTGCGCAAAAGTTTTACATATCGAGAGAACACACAGAGGAGAAACCAACAAACAA 1560  
Qy 1798 GAACAAAGATATCAATGGCACTTGATCTTAACTTCAAGCTCAGTTAAACACCGATTGAT 1857  
Db 1561 GAACAAAGATATCAATGGCACTTGATCTTAACTTCAAGCTCAGTTAAACACCGATTGAT 1620  
Qy 1858 GATCAAGAGAGAGAGAGAACACAGGATTTCTTGGATCGATTAGATGCTTCAAGCTA 1917  
Db 1621 GATCAAGAGAGAGAGAGAACACAGGATTTCTTGGATCGATTAGATGCTTCAAGCTA 1680  
Qy 1918 ATCAGTAGAGGAGAGAACAGGTTTTTAAACCATACAAAGATGTTCCATGGAAGCCAAAGAA 1977  
Db 1681 ATCAGTAGAGGAGAGAACAGGTTTTTAAACCATACAAAGATGTTCCATGGAAGCCAAAGAA 1740  
Qy 1978 AGTAGAATCTTCAACAACTATCTATCATGTTGGAACAGAAAGATCCCAACCGGATG 2037  
Db 1741 AGTAGAATCTTCAACAACTATCTATCATGTTGGAACAGAAAGATCCCAACCGGATG 1800  
Qy 2038 CGGTTGGAAGTCAAGCTTCCACATGA 2064  
Db 1801 CGGTTGGAAGTCAAGCTTCCACATGA 1827

## RESULT 4

US-10-084-553-1  
; Sequence 1, Application US/10084553  
; Publication No. US20030056247A1

## ; GENERAL INFORMATION:

; APPLICANT: Tobin, Elaine  
; APPLICANT: Wang, Zhi-Yong  
; APPLICANT: Sun, Lin  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Phytocrome Regulated Transcription Factor for Control  
; TITLE OF INVENTION: of Higher Plant Development  
; FILE REFERENCE: 023070-124200US  
; CURRENT APPLICATION NUMBER: US/10/084,553  
; PRIOR FILING DATE: 2002-07-09  
; PRIOR APPLICATION NUMBER: US 08/843,572  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4344  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

; FEATURE:  
; OTHER INFORMATION: phytocrome-regulated transcriptional factor CCA1  
; OTHER INFORMATION: genomic clone

; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1332)..(1370)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1371)..(1448)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1449)..(1560)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1561)..(1648)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1649)..(1710)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1711)..(2189)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (2190)..(2384)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (2385)..(2471)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (2472)..(2645)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (2646)..(2728)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (2729)..(3610)  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (3611)..(3680)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (3681)..(4043)  
; US-10-084-553-1

Query Match 67.88; Score 1527.6; DB 14; Length 4344;  
Best Local Similarity 87.9%; Pred. No. 0;  
Matches 1803; Conservative 0; Mismatches 9; Indels 240; Gaps 3;

Qy 441 TTTCTCCAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGTAGCTATGGTCAAGCGCTAGA 500  
Db 2180 TCTCTCTCAGGTAGAGAAAGAGGCTGAAGCTAAAGGTGTAGCTATGGTCAAGCGCTAGA 2239  
Qy 501 CATAGCTATTCTCTCCAGCGGCTAAGCGTAAACCAAAACAATCTTATCTCGAAAGAC 560  
Db 2240 CATAGCTATTCTCTCCAGCGGCTAAGCGTAAACCAAAACAATCTTATCTCGAAAGAC 2299  
Qy 561 GGGAAGTGAACCATCTTATGTCAAAACGGGTGTGAATGATGGAAGAGTCCCTTGG 620  
Db 2300 GGGAAGTGAACCATCTTATGTCAAAACGGGTGTGAATGATGGAAGAGTCCCTTGG 2359  
Qy 621 ATCAGAAAAAGTGTCCGATCTCTGCTGCTGATGTTTTCATGGTTCATATGGCATCTTTTGGCAGT 643  
Db 2360 ATCAGAAAAAGTGTCCGATCTCTGCTGCTGATGTTTTCATGGTTCATATGGCATCTTTTGGCAGT 2419  
Qy 644 -----AGATGGCCAA 653  
Db 2420 GTGTCACTTGTCTCTCATGTTTATTAATACAGATTGTGCTTCTGTTTATAGATGGCCAA 2479  
Qy 654 TGAAGATCGACACAACTCAAGCCTCAAGAGAAACTCTCGAGAGACAACTGTTTCAGA 713  
Db 2480 TGAAGATCGACACAACTCAAGCCTCAAGAGAAACTCTCGAGAGACAACTGTTTCAGA 2539  
Qy 714 TTGTTTCACTCATCAGTATCTCTCTGCTGATCTCTCCATGCAATAAAAAGTTGTATAGAGAC 773

Db 2540 TTGTTTCACTCATCAGTATCTCTGCTGCATCTCCATGAATAAAGTTGTATAGAGAC 2599  
Qy 774 ATCAAAACGACGACATTTCCGCGAGTTCTTCCGCTTCCAGGAAAG----- 817  
Db 2600 ATCAAAACGACGACATTTCCGCGAGTTCTTCCGCTTCCAGGAAAGGTTAAACAAATCT 2659  
Qy 818 ----- 817  
Db 2660 TTCATTGCTATTTGAGGTTTTAAGCGATTAGTACTTTTCATGAACCTAAACCGTGGG 2719  
Qy 818 -----AGGAAGTCAGATAACAGGGTAAGAAAGGAGTCAAACTCAGATTTCGAATGCA 870  
Db 2720 GAATAACAGGGAAGTCAGATAACAGGGTAAGAAAGGAGTCAAACTCAGATTTCGAATGCA 2779  
Qy 871 AAATCTCTGGAACCGGTAATGACCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG 930  
Db 2780 AAATCTCTGGAACCGGTAATGACCAAGGACCTCAGACTTATCCGATGCATATCCCTGTG 2839  
Qy 931 CTAGTGCCATTTGGGAGCTCAATAACAAAGTTCTCTATCACATCCTCCTTCAGAGCCAGAT 990  
Db 2840 CTAGTGCCATTTGGGAGCTCAATAACAAAGTTCTCTATCACATCCTCCTTCAGAGCCAGAT 2899  
Qy 991 AGTCATCCCAACACAGTTGACGAGATATCAGTCGTTTCTTAATCATATAATGTCAACC 1050  
Db 2900 AGTCATCCCAACACAGTTGACGAGATATCAGTCGTTTCTTAATCATATAATGTCAACC 2959  
Qy 1051 CTTTTCACAAACACCGGCTCTTTATCTACTGCGCAACTTTCCGCTCATCATTTTGGCCCTCC 1110  
Db 2960 CTTTTCACAAACACCGGCTCTTTATCTACTGCGCAACTTTTGGCCCTCATCATTTTGGCCCTCC 3019  
Qy 1111 GATTCTAGTGGTGCTCACCTGTTTCCAGGAACTCACCTCCGAATCTGGCTGCCATGGCC 1170  
Db 3020 GATTCTAGTGGTGCTCACCTGTTTCCAGGAACTCACCTCCGAATCTGGCTGCCATGGCC 3079  
Qy 1171 GCAGCCACTGTTGCAGCTGCTAGTGTCTTGGTGCTGCCAATGATTAATACCTTTATGT 1230  
Db 3080 GCAGCCACTGTTGCAGCTGCTAGTGTCTTGGTGCTGCCAATGATTAATACCTTTATGT 3139  
Qy 1231 GCTCCTCTTAGTTAGTTCAGTGGTTTCACTAGTCATCTCCTCATCTACTTTTGGACCATCATGT 1290  
Db 3140 GCTCCTCTTAGTTAGTTCAGTGGTTTCACTAGTCATCTCCTCATCTACTTTTGGACCATCATGT 3199  
Qy 1291 GATGTAGATACAAAAAGCAAGCATTTTACAACATGGTTCTGTGCAGAGCCGAGAGCAA 1350  
Db 3200 GATGTAGATACAAAAAGCAAGCATTTTACAACATGGTTCTGTGCAGAGCCGAGAGCAA 3259  
Qy 1351 GAACACTCCGAGGATCAAAAGGCTCGATCTTCACTGCACTCAGAGGATGTTGAAAATAAG 1410  
Db 3260 GAACACTCCGAGGATCAAAAGGCTCGATCTTCACTGCACTCAGAGGATGTTGAAAATAAG 3319  
Qy 1411 AGTAAACCAAGTTTGTATGAGCAGCCCTTCTGCAACACTCAGAGTGTGATGCAAGGGTTCA 1470  
Db 3320 AGTAAACCAAGTTTGTATGAGCAGCCCTTCTGCAACACTCAGAGTGTGATGCAAGGGTTCA 3379  
Qy 1471 GATGGAGCAGAGACAGAAAAAAGTTGACCGGTCCTCGTGTGGCTCAAAACACTCCGTCG 1530  
Db 3380 GATGGAGCAGAGACAGAAAAAAGTTGACCGGTCCTCGTGTGGCTCAAAACACTCCGTCG 3439  
Qy 1531 AGTAGTGATCATGTTGAGGCGGATGTCATCAGAAGGCAAGGATGCAACCAATGTGTAG 1590  
Db 3440 AGTAGTGATCATGTTGAGGCGGATGTCATCAGAAGGCAAGGATGCAACCAATGTGTAG 3499  
Qy 1591 GTGAAAGAAACGAATGAAGACACTTAATAAACCTCAAACTTCAGAGTCCAATGCACGCCGC 1650  
Db 3500 GTGAAAGAAACGAATGAAGACACTTAATAAACCTCAAACTTCAGAGTCCAATGCACGCCGC 3559  
Qy 1651 AGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGTGCTGACGA----- 1700  
Db 3560 AGTAGAATCAGTCCCAATATAACCGATCCATGGAAGTCTGTGTGCTGACGAGTACTTACT 3619  
Qy 1701 ----- 1700

Db 3620 TGGACTAAAGATCAACTTCTCTTTATTTCAAATCATTTTCTCATATAAATATTGTACATTC 3679  
Qy 1701 GGGTCGAATTCGCTTCCAAGCTCTCTTCTCCAGAGAGGTATTGGCGCAAGTTTTTACATA 1760  
Db 3680 GGGTCGAATTCGCTTCCAAGCTCTCTTCTCCAGAGAGGTATTGGCGCAAGTTTTTACATA 3739  
Qy 1761 TCGAAGAAGAACACAGAGAGGAAGAACAAACAACAAGAACAAAGATATCCAAATGGCACT 1820  
Db 3740 TCGAAGAAGAACACAGAGAGGAAGAACAAACAACAAGAACAAAGATATCCAAATGGCACT 3799  
Qy 1821 TGATCTTAACTTTCACAGCTCAGTTTAAACACACAGTTTGTATCAAGAGGAGAGAGAAACAC 1880  
Db 3800 TGATCTTAACTTTCACAGCTCAGTTTAAACACACAGTTTGTATCAAGAGGAGAGAGAAACAC 3859  
Qy 1881 AGGATTTCTTGGATTCGATAGATGCTTCAAGCTTAAGTGTAGAGGAGAGAAACAGGTTT 1940  
Db 3860 AGGATTTCTTGGATTCGATAGATGCTTCAAGCTTAAGTGTAGAGGAGAGAAACAGGTTT 3919  
Qy 1941 TAAACCATACAAAGATGTTTCCATGGAGGCCAAAGAAAGTAGAATCCTCAACAACAATCC 2000  
Db 3920 TAAACCATACAAAGATGTTTCCATGGAGGCCAAAGAAAGTAGAATCCTCAACAACAATCC 3979  
Qy 2001 TATCATTCATGTGGAACAGAAAGATCCAAACCGATCGGTTGGAAACTCAAGCTTCCAC 2060  
Db 3980 TATCATTCATGTGGAACAGAAAGATCCAAACCGATCGGTTGGAAACTCAAGCTTCCAC 4039  
Qy 2061 ATGAGACTCTATTTTCTATCTGATCTGTTTGTGTACTCTGTTTAAAGTTTTCAGAGCA 2120  
Db 4040 ATGAGACTCTATTTTCTATCTGATCTGTTTGTGTACTCTGTTTAAAGTTTTCAGAGCA 4099  
Qy 2121 CTGCTACATTTCTTTTCTTTTGGAGCCCTTGTATTTGTTCTGTGTCATAGTCTTCC 2180  
Db 4100 CTGCTACATTTCTTTTCTTTTGGAGCCCTTGTATTTGTTCTGTGTCATAGTCTTCC 4159  
Qy 2181 TGTAACATTTGACTCTGATTTATTCACAAATCATAAACTGTTTTAAATCTTTTTTTTCCA 2240  
Db 4160 TGTAACATTTGACTCTGATTTATTCACAAATCATAAACTGTTTTAAATCTTTTTTTTCCA 4219  
Qy 2241 AAAAAAAAAA 2252  
Db 4220 ACCTGGAAAGAA 4231

## RESULT 5

US-09-770-961-961  
; Sequence 961, Application US/09770961  
; Publication No. US20030115639A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kricker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE REFERENCE: 2026 (PARA-01SPRV)  
; CURRENT APPLICATION NUMBER: US/09/770,961  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,466

; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 961  
; LENGTH: 479  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-961-961

Query Match 21.3%; Score 479; DB 11; Length 479;  
Best Local Similarity 100.0%; Pred. No. 6.6e-124; Indels 0; Gaps 0;  
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 ATAGCTATTCCTCTCCACGGCTTAAGCGTAACCAACCAATCTTATCTCTGAAAGACG 561  
Db 1 ATAGCTATTCCTCTCCACGGCTTAAGCGTAACCAACCAATCTTATCTCTGAAAGACG 60  
QY 562 GGAAGTGGACGATCCTTATGTCAAAACCGGTGTGAATGATGAAAGAGTCCCTTGA 621  
Db 61 GGAAGTGGACGATCCTTATGTCAAAACCGGTGTGAATGATGAAAGAGTCCCTTGA 120  
QY 622 TCAGAAAAAGTGTGCGATCTCGATGCGCAATGAAGATCGACAACCAATCAAGCCCTGAA 681  
Db 121 TCAGAAAAAGTGTGCGATCTCGATGCGCAATGAAGATCGACAACCAATCAAGCCCTGAA 180  
QY 682 GAGAAAACTCTGAGGAAGACAACTGTTGAGATTTGTTCACTCATCAGTATCTCTGCT 741  
Db 181 GAGAAAACTCTGAGGAAGACAACTGTTGAGATTTGTTCACTCATCAGTATCTCTGCT 240  
QY 742 GCATCTCCATGAATTAAGTTCTATAGACATCAACCGCAAGCACTTTCGGGAGTTC 801  
Db 241 GCATCTCCATGAATTAAGTTCTATAGACATCAACCGCAAGCACTTTCGGGAGTTC 300  
QY 802 TTGCTCTTCAGGAAGGGAAGTCAAGATAACAGGGTAAAGAAAGGAGTCAAACTCAGAT 861  
Db 301 TTGCTCTTCAGGAAGGGAAGTCAAGATAACAGGGTAAAGAAAGGAGTCAAACTCAGAT 360  
QY 862 TTGAATGCAAAATCTCTGAAAAACGGTAATGAGCAAGGACCTCAGACTTATCCGATGCAT 921  
Db 361 TTGAATGCAAAATCTCTGAAAAACGGTAATGAGCAAGGACCTCAGACTTATCCGATGCAT 420  
QY 922 ATCCCTGTCTAGTGCATTTGGGAGTCAATGAAGTTCTTATCAGATCTCTCTTC 980  
Db 421 ATCCCTGTCTAGTGCATTTGGGAGTCAATGAAGTTCTTATCAGATCTCTCTTC 479

## RESULT 6

US-09-934-455-119  
; Sequence 119, Application US/09934455  
; Publication No. US20030121070A1  
; GENERAL INFORMATION:  
; APPLICANT: Adam, Luc  
; APPLICANT: Creelman, Robert  
; APPLICANT: Dubell, Arnold  
; APPLICANT: Heard, Jacqueline  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Keddle, James  
; APPLICANT: Pilgrim, Marsha  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Reuber, Lynne  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Pineda, Omalra  
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV  
; FILE REFERENCE: MBI-0025  
; CURRENT APPLICATION NUMBER: US/09/934,455  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/227439  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: MBI-0022  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: MBI-0023  
; PRIOR FILING DATE: 2001-04-17

; NUMBER OF SEQ ID NOS: 516  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 119  
; LENGTH: 2526  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; NAME/KEY: CDS  
; LOCATION: (338)..(2275)  
; OTHER INFORMATION: G680  
US-09-934-455-119

Query Match 9.6%; Score 217.2; DB 11; Length 2526;  
Best Local Similarity 49.9%; Pred. No. 1.1e-49;  
Matches 986; Conservative 0; Mismatches 823; Indels 165; Gaps 11;

QY 235 GTGATGGAGACAATTCGTCTGGAGAAAGATCTCGTTATTAAAGATCTCGAAGCCATATACG 294  
Db 335 GTTATGGATTAATAATACATCTGGAGAAAGATTTATTAGCTAAGGCAAGAACCCATATACA 394  
QY 295 ATACAAAGCAACGTGAAAGGTGACTGAGGAAGACATAATAGATTTCATTGAAGCTTTG 354  
Db 395 ATACAAAGCAGCGAGAGCGATGACTTGAGGATGAGCATGAGAGTTTCTAGAAAGCTTTG 454  
QY 355 AGGCTTTATGGTAGAGCATGGCAGAAAGATTGAAGAACATGTAGCAACAAAAAATGCTGTC 414  
Db 455 AGGCTTTATGGAGAGCTTGGCAACGAATTTGAAGAACATATTGGGACAAAGACTGCTGTT 514  
QY 415 CAGATAAGAGTCACGCTCAGAAATTTTCTCAAGGTAGAGAAAGAGGCTGAAGCTAAA 474  
Db 515 CAGATCAGAAGTCATGCACAAAAGTTCTTCAAAAAGTTGGAGAAAGAGGCTGAAGTTAAA 574  
QY 475 GGTGTAGCTATGGTCAAGCGCTAGACATAGCTATTCTCTCCACGGCTTAAGCGTAAA 534  
Db 575 GGCATCCCTGTTGCCAAGCTTTGGACATAGAAATTCGGCTCTCTCTCTTAAACGAAA 634  
QY 535 CCAAAACAATCCTTTATCTCGAAAAGACGGGAAGTGAACGATCCTTATGTCAAAAACGGGT 594  
Db 635 CCCAATACCTCTTATCTCGAAAACCTGGGAACACACGGTACATCTTCTCTCAAGTATCA 694  
QY 595 GTGAATGATGGAAGAGTCCCTTGGATCAGAAAAAGTGTGCGATCTCTGAGATGCCCAAT 654  
Db 695 TCAGCAAAAGATGCAAAAACCTTGTTCATCGGCTCTTCTTCAAGTTGAATCAGGCGTTC 754  
QY 655 GAAGATCGACAACAATCAAAAGCCTGAAGACAAAC-----TCTGCAGAA 699  
Db 755 TTGATTTGGAAAAAATGCCGTTCTCTGAAAAACATCAACTGGAAGAAAAAATCAAGAT 814  
QY 700 GACAACCTGTTTCAGATTGTTTCACTCATCATATCTCTCTGCTGCTCATCTCCATGAATAAA 759  
Db 815 GAGAATTCGTCGGGTGTTTCTACTGTGAACAAGTATCCCTTACCACGAAACAGGTAAAT 874  
QY 760 AGTTGTATAGAGACATCAACGCAAGCACTTT-----CGCGAGTTCTTTCCTTC 809  
Db 875 GCGCATTGAAACAAGTAAAGACCTCAACTGTGGACAACCGGTTCAAGATGTTCCCAAG 934  
QY 810 ACGGGAAGAGGGAAGTCAAGATAACAGGTA-----AGAAAGGAGTCAAC 855  
Db 935 AAGAACAAAGACAAAGATGGTAACGATGGTACTACTGTGCACAGCATGCAAAACTACCC 994  
QY 856 TCAGATTGGAATGCAAAAATCTCTGAAAAACGGTAATGAGCAAGGACCTCAGACTTATCC- 914  
Db 995 TGGCAATTCACGCGCAGATATTGTGAACGGGAATATAGCAAAATGCCCTCAAAATCATCCC 1054  
QY 915 -----G 915  
Db 1055 TCAGGTATGGTATCTCAAGACTTTCATGTTTTCATCTCTATGAGAGAAACTCAGGGCAC 1114  
QY 916 ATGCATATCCCTGTGCTAGTGCATTTGGGAGCTCAATAACAAGTTCTCTATCACATCCT 975  
Db 1115 GCAATCTTCAAGCTACAAAGCATCTGCTACTACTACAGTTCTCATCAAGGTTTCCA 1174  
QY 976 CCTTCAGAGCCAGATAGTATCCCCACACAGTTTGCAGGAGATTATCAGTCTGTTCTCTAAT 1035



Db 815 GAGAAATGTCGGGTGTTTCTACTGTGAACAAGTATCCCTTACCAACGAAACAGGTAAGT 874  
QY 760 AGTTGTATAGACATCAACGCAAGCACTTT-----CCGCGAGTTCTTGCGCTTC 809  
Db 875 GCGGACATTTGAACAAGTAAGACCTCACTGTGGACAACCGGTTCAAGATGTTCCCAAG 934  
QY 810 ACGGGAAGGGAAGTFCAGATAACAGGGTA-----AGAAAGGAGTCAAAAC 855  
Db 935 AAGAACAAAGACAAGATGGTAACGATGGTACTACTGTGCACAGCATGCAAAACTACCCT 994  
QY 856 TCAGATTTGAATGCAAAATCTCTGGAAGAGGTAAATGACAGGACCTCAGACTTATCC- 914  
Db 995 TGGCATTTCCAGCGCATATTGTGAACGGGAATATAGCAAAATGCCCTCAAAATCATCCC 1054  
QY 915 -----G 915  
Db 1055 TCAGGTATGGTATCTCAAGACTTCATGTTTCATCCTATGAGAGAAGAACTACGGGCAAC 1114  
QY 916 ATGCATATCCCTGTGCTAGTGCCATTGGGAGCTCAATAACAAGTTCTCTATCACATCCT 975  
Db 1115 GCAAAATCTTCAAGCTACAACAGCATCTGCTACTACTACAGCTTCTCATCAAGCGTTTCCA 1174  
QY 976 CTTTCAGAGCCAGATAGTATCCCAACACAGTTGCGAGGAGATATCAGTCGTTTCTTAAT 1035  
Db 1175 GCTTGTCTTCACAGGATGATTACCGTTCTTCTCCAGATATCATCTACTTTCTCCAAT 1234  
QY 1036 CATATAATGCAACCCCTTTTCAAAACACCGGCTCTTTATACTGCCGCAACTTTTCGGCTCA 1095  
Db 1235 CTTATTATGTCAACTCTCTACAGATCTCTGAGCTCATGCTGCAGCTACATCTCGTGT 1294  
QY 1096 TCATTTTGGCCTCCCG-----ATTCTAGTGGTGGCTCACCTGTTCCAGGGAAC 1143  
Db 1295 TCGGTCTGGCCTTATCGAGTGTGCGGAATCTGCTGATTCTATCAACCCCAATGAGCTCT 1354  
QY 1144 TCACCTCCGAATCTGGCTGCCATGCCGAGCACTGTTGCGAGCTGCTAGTCTGGTGG 1203  
Db 1355 TCTCTCCAAGTATACTGCCATTTGCCGTGTACAGTAGCTGCTGCAACTGCTGGTGG 1414  
QY 1204 GCTGCCAATGGATTATTAACCTTTATGTGCTCTCTTAGTTTCAAGTGGTTTCACTAGTAT 1263  
Db 1415 GCTTCTCATGGACTTCTCTGTATGCTGCTCCAGCTCCATACATGTTGTTCAATTCTCA 1474  
QY 1264 CTTCCATCTACTTTTGGACATCATGTGATGATGATGATGATGATGATGATGATGATGAT 1323  
Db 1475 ACTGTTGCAAGTTTCCAACTCCAGCAATGACTGAAATGGATACCGTTGAAATACTCAAC 1534  
QY 1324 CATGGTCTGTGCAGAGCCGAGAGCAAGCACTCCGAGGCATCAAGGCTCGATCTTCA 1383  
Db 1535 TTTGAGAAAACAAACACAGCTCTGCAAGATCAAACTTGGCTTCGAAATCTCCAGCTTCA 1594  
QY 1384 CTGGACTCAGAGGATGTTGAAAAT-----AAGAGTAAACCACTGTTGTCATGAGCAG 1434  
Db 1595 TCATCTGATGATTTCAGATCAGACTGGAGTAAACCAAGTAAATGCCGACTCAAAACCAAT 1654  
QY 1435 CTTCTGCAACCTGAGAGTGAATGCAA-----AGGTTTCAGATGAGGAGCAGGA 1482  
Db 1655 GATGATAAAATTCAGAGGAGTGTGTTTACTGCGCGCTGTGCATGACTCAAAACACTGCC 1714  
QY 1483 GACAGAAAACAAAGTTGACCGGTCCTGCTGCTGCTCAAACTCCGTCGAGTAGTATGAT 1542  
Db 1715 AAGAAAAATCTTTGGGACCGCTCATGCTGCTGCTCAAAATACACCTTTCAGGAGTAGC 1774  
QY 1543 GTTAGGCGGATGCATCAAAAAGGCAAGAGGATGGCAACCAATGGTGGAGTGAAGAAACG 1602  
Db 1775 GAACTGATGCATTAGATAAAATGGAGAAAGATTAAGAGGATGTGAGGAGACAGATGAG 1834  
QY 1603 AAT-----GAAGACACTTAATAACCTCAAACTTCAGAGTCCAAATGACGCGCGCAGTAGA 1656  
Db 1835 AATCAGCCAGATGTTTATTGAGTTAAATAACCGTAAAGATTAAATGAGAGCAACCAACAG 1894  
QY 1657 ATCAGCTCCAATATACCGATCCATGGAGTCTGTGCTGACGAGGTCGAATTCGCTTC 1716

Db 1895 AACAAACAATGCAACTACTGATTCTGCGAAAGGAAGTCTCCGAAGAGGGTCTGTATAGCGTTT 1954  
QY 1717 CAAGCTCTCTTCCAGAGAGGTATTGCCGCAAGTTTTACATATCGAGAAGAACACAGA 1776  
Db 1955 CAGGCTCTCTTTGCAAGAGAAAGATTGCCCTCAAAGCTTTTCGGCTCTCTCAAGTGGCAGAG 2014  
QY 1777 GAGGAGAAACAACAACAACAAGATATCCAAATGGCACTTGTATCTTTAACTTCCACA 1836  
Db 2015 AATGTGAATAGAAAAACAAAGTGACACGT-----CAATGCCATTGGCTCCTTAATTTCAA 2068  
QY 1837 GCTCAGTTAAACACCAGTTGATGATCAAGAGGAGAGAAACACAGAGATTTCTTGGAAATC 1896  
Db 2069 AGCCAGGATTCTTGTGCTGCAGACCAAGA-----AGGAGTAGTAATGATC 2113  
QY 1897 GGATTAGATGCTTCAAGCTAATGACTAGAGGAAGAACAGGTTTAAACCATACAAAGA 1956  
Db 2114 GGTGTTGGAACATGCAAGAGTCTTAAACGAGACACAGGATTTAAGCCATACAAAGA 2173  
QY 1957 TGTTCATGGAAGCCAAAGAAAGTAGAATCTCTCAACAAACAAATCTTATCTATGTGAA 2016  
Db 2174 TGTTCATGGAAGTGAAGAGAGAGCCAAAGTTGGGAACATAAA-----CAATCAAAGTAT 2227  
QY 2017 CAGAAAGATCCAAACCGGATGCGGTTGGAAACTCAAGCTTCCACATGAGACTCT 2070  
Db 2228 GAAAAAGTCTGCAAAAGGCTTCGATTGGAAGGAGAGAGCTTCTACATGACAGACT 2281

## RESULT 8

US-09-938-842A-1116  
; Sequence 1116, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPT1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1116  
; LENGTH: 1938  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1116

Query Match 9.4%; Score 213; DB 10; Length 1938;  
Best Local Similarity 49.9%; Pred. No. 1.4e-48;  
Matches 980; Conservative 0; Mismatches 820; Indels 165; Gaps 11;

QY 238 ATGGAGACAAATTCGTCTCGAGAAGATCTGGTTATTAAAGACTCGGAAGCCATATACGATA 297  
Db 1 ATGGGATACTAATACATCTCGAGAAGAAATATTAGCTAAGCAAGAAAGCCATATACAATA 60  
QY 298 ACAAGACACGTAAGGTTGAGTGGAGGAAACATATAGATTCTATTGAGCTTTGAGG 357  
Db 61 ACAAGACGAGAGCGATGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 120  
QY 358 CTTTATGTTAGAGCTGCGAGAGATTGAAGAACATGTAGCAACAAAACTGCTGTCCAG 417  
Db 121 CTTTATGGAAGAGCTTGGCAACGAATTGAAGAACATATTGGGACAAAGACTGCTGTTCAG 180  
QY 418 ATGAAGATCAGCTTCAGAAATTTTCTCAGAGGTAGAGAAAGAGGCTGAAGCTAAAGGT 477  
Db 181 ATCAGAAGTCATGCACAAAAGTTCTTCAAAAAGTTGGAGAAAGAGGCTGAAGTTAAAGGC 240







```
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Zhang, James
; APPLICANT: Benito, Maria-Ines
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Fromm, Mike
; TITLE OF INVENTION: PLANT GENE SEQUENCES I
; FILE REFERENCE: MBI-0003
; CURRENT APPLICATION NUMBER: US/10/295,403
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 60/101,349
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/103,312
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/108,734
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/113,409
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 1237
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)..(951)
; OTHER INFORMATION: G673
US-10-295-403-49

Query Match      3.7%; Score 83; DB 14; Length 1237;
Best Local Similarity 60.4%; Pred. No. 4e-12;
Matches 137; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 237 GATGGAGACAAATTCGTCTCGAGAGATCTGGTTATTAACTCGGAAGCCATACGAT 296
Db 144 GATTGCACACCGGAGCTGGTGAACACCGGAGAGAGAGGTGAGGAAGCTTATACAAT 203
QY 297 AACAAAGCAACGTGAAGGTGGACTGAGGAGAACATATAGATTTCATTGAAGCTTTGAG 356
Db 204 CACCAAGCTAGAGAGAGTTGGACTGAAGGAGAACACGACAGTTTCTGGAAGCTCTTCA 263
QY 357 GCTTTATGTAGACATGCGCAGAGATTGAAGAACATGTAGCAACAAACTGCTGTCCA 416
Db 264 ATTGTTTGTGCTGACTGGAAGAAAGATAGAAAGATTTTGTGGTTCAAGACAGTTATTCA 323
QY 417 GATAAGAAGTCACGCTCAGAAATTTTCTCCAAGGTAGAGAAAGAGG 463
Db 324 GATCAGAGGACCATGCCCAAAATACCTTCTAAAGGTCCAAAATAATG 370

RESULT 13
US-09-878-574-707
; Sequence 707, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775

; SEQ ID NO 707
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LfB3028-047-Q1-B1-E4
US-09-878-574-707

Query Match      3.3%; Score 74.8; DB 10; Length 365;
Best Local Similarity 65.7%; Pred. No. 3.3e-10;
Matches 109; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 274 AAGACTCGGAAGCCATATACGATAACAACAAACGTCGAAAGGTGGACTGAGGAAGAACAT 333
Db 198 AAGTAAGAAACCTTTAGACCATAACCAAGTCCAGGGAGAGTTGGACTGAGGAAGACAC 257
QY 334 AATGATTTCATTGAAGCTTTGAGGCTTTATGTTAGACCATGGCAGAGATTGAAGAACAT 393
Db 258 GACAAGTTTCTCGAAGCTCTTCAATTATTGACAGGGACTGGAAGAAAATTGAAGATTTT 317
QY 394 GTAGCAACAAAACTGCTGCCAGATAAGAAAGTCACGCTCAGAAAT 439
Db 318 GTAGGTTCAAAAAACAGTTATTTCAGATTTCAGCCATGCTCAGAAAT 363

RESULT 14
US-09-814-353-4739
; Sequence 4739, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4739
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 55, 54, 159, 161, 174, 177, 178, 180, 182, 185, 212, 215,
; LOCATION: 216, 219, 220, 226, 237, 247, 248, 265, 269, 278, 281, 283,
; LOCATION: 289, 294, 298, 299, 300, 301, 302, 304, 305, 307, 308, 316,
; LOCATION: 327, 333, 334, 335, 336, 339, 343, 344, 345, 348, 368
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 377, 383, 386, 388, 390, 409, 429, 431, 433, 439, 447, 450,
; LOCATION: 452, 460, 462, 473, 474, 475, 477, 484, 486, 508, 509, 510,
; LOCATION: 511, 513, 514, 515, 518, 519, 520, 522, 525, 528, 530, 532,
; LOCATION: 533, 541, 542, 543, 544, 547, 550, 552, 561, 564, 570
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 573, 575, 578, 584, 600, 601, 603, 606, 607, 616, 623, 625,
```

